

**UNIVERSIDADE FEDERAL DE SANTA CATARINA
DEPARTAMENTO DE INFORMÁTICA E ESTATÍSTICA**

Rodrigo Tridapalli Fóes Linhares

**MAPEAMENTO DA RETINA A PARTIR DE IMAGENS
DE UMA LÂMPADA DE FENDA USANDO TÉCNICAS
DE AJUSTE GLOBAL**

Florianópolis

2016

Rodrigo Tridapalli Fóes Linhares

**MAPEAMENTO DA RETINA A PARTIR DE IMAGENS
DE UMA LÂMPADA DE FENDA USANDO TÉCNICAS
DE AJUSTE GLOBAL**

Dissertação submetida ao Programa
de Pós-graduação em Ciência da Com-
putação para a obtenção do Grau de
Mestrado em Ciência da Computação.

Orientador

Universidade Federal de Santa Cata-
rina: Prof. Dr. Ricardo Alexandre
Reinaldo de Moraes

Coorientador

Universidade Federal de Santa Cata-
rina: Prof. Dr. Rogério de Almeida
Richa

Florianópolis

2016

Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Linhares, Rodrigo Tridapalli Fôes

Mapeamento da retina a partir de imagens de uma lâmpada de fenda usando técnicas de ajuste global / Rodrigo Tridapalli Fôes Linhares ; orientador, Ricardo Alexandre Reinaldo de Moraes ; coorientador, Rogério de Almeida Richa. - Florianópolis, SC, 2016.
66 p.

Dissertação (mestrado) - Universidade Federal de Santa Catarina, Centro Tecnológico. Programa de Pós-Graduação em Ciência da Computação.

Inclui referências

1. Ciência da Computação. 2. Processamento de Imagens. 3. Ajuste Global. 4. Lâmpada de Fenda. 5. Oftalmologia. I. de Moraes, Ricardo Alexandre Reinaldo. II. Richa, Rogério de Almeida. III. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Ciência da Computação. IV. Título.

Rodrigo Tridapalli Fóes Linhares

**MAPEAMENTO DA RETINA A PARTIR DE IMAGENS
DE UMA LÂMPADA DE FENDA USANDO TÉCNICAS
DE AJUSTE GLOBAL**

Esta Dissertação foi julgada aprovada para a obtenção do Título de “Mestrado em Ciência da Computação”, e aprovada em sua forma final pelo Programa de Pós-graduação em Ciência da Computação.

Florianópolis, 01 de Agosto 2016.

Prof. Dr. Carina Friedrich Dorneles
Coordenadora
Universidade Federal de Santa Catarina

Banca Examinadora:

Prof. Dr. Ricardo Alexandre Reinaldo de Moraes
Orientador
Universidade Federal de Santa Catarina

Prof. Dr. Rogério de Almeida Richa
Coorientador
Universidade Federal de Santa Catarina

Prof. Dr. Patricia Della Mea Plentz
Universidade Federal de Santa Catarina

Prof. Dr. Aldo von Wangenheim
Universidade Federal de Santa Catarina

Prof. Dr. Agma Juci Machado Traina
Universidade de São Paulo

RESUMO

Lâmpadas de fenda são ferramentas comumente usadas para auxiliar cirurgias minimamente invasivas da retina, com o objetivo de diagnosticar e tratar doenças que afetam a visão. Na literatura, os métodos propostos para auxílio computacional em cirurgias feitas com lâmpadas de fenda sofrem de falhas causadas por imperfeições no registro de imagens e acúmulo de erros. Essa dissertação propõe um algoritmo de ajuste fino de mapas da retina, com o objetivo de alinhar mosaicos construídos a partir de imagens obtidas pela lâmpada de fenda, e melhorar a qualidade do seu produto final, aumentando assim o potencial de auxílio em cirurgias aplicadas à retina. A estratégia proposta é um método direto que alinha o conjunto de imagens que compõe o mosaico simultaneamente através da minimização de uma medida global de erro e do uso de um modelo de transformação não-rígido. Os resultados obtidos mostram uma melhora na qualidade dos mosaicos, além da correção de borrões e outros erros de alinhamento.

Palavras-chave: Processamento de imagens. Lâmpada de fenda. Oftalmologia.

ABSTRACT

Slit lamps are tools commonly used to aid in minimally invasive surgeries on the retina, with the goal of diagnosing and treating diseases that affect vision. In the literature, the methods proposed for computational assistance in procedures done with slit lamps suffer from issues caused by imperfections in image registration and error accumulation. This work proposes a fine adjustment algorithm for retina maps, which seeks to align mosaics built from images obtained with a slit lamp, and improve the quality of the final result, thus increasing the potential for computational assistance provided to surgeries performed on the retina. The proposed strategy is a direct method that simultaneously aligns all images that make up a retina map through the minimization of a global error measure and the use of a non-rigid transformation model. Results show an improvement in overall map quality, as well as the elimination of blurs and other alignment errors.

Keywords: Image processing. Slit lamp. Ophthalmology.

SUMÁRIO

1	INTRODUÇÃO	11
1.1	CONTEXTUALIZAÇÃO	11
1.1.1	Desafios	13
1.2	OBJETIVO GERAL	16
1.2.1	Objetivos Específicos	16
1.3	CONTRIBUIÇÕES	16
2	REVISÃO BIBLIOGRÁFICA	19
2.1	MAPEAMENTO DE IMAGENS MÉDICAS	20
2.1.1	Métodos Baseados em Pontos de Interesse	20
2.1.2	Métodos Diretos	30
3	METODOLOGIA	37
3.1	MAPEAMENTO DA RETINA	39
3.1.1	Pré-processamento	39
3.1.2	Rastreamento	40
3.1.3	Seleção Eficiente de Pixels	43
3.2	AJUSTE GLOBAL	44
3.2.1	Pré-processamento	45
3.2.2	Ajuste	45
3.3	RESULTADOS	51
4	CONCLUSÃO	57
	REFERÊNCIAS	59

1 INTRODUÇÃO

1.1 CONTEXTUALIZAÇÃO

Doenças que afetam a retina, como a retinopatia diabética e a degeneração macular relacionada à idade, estão entre as principais causas de danos à visão, sendo responsáveis por 6% dos casos mundiais de cegueira (World Health Organization, 2010). Essas patologias causam grandes áreas de isquemia (regiões nas quais o fluxo sanguíneo é debilitado), ocasionando irregularidades na formação de novos vasos sanguíneos. Sem tratamento, os sintomas pioram progressivamente, começando com embaçamento e escurecimento da visão e sensibilidade à luz, podendo resultar em descolamento da retina ou em cegueira total.

As regiões de isquemia são comumente tratadas através da fotocoagulação a laser, que consiste na cauterização das áreas afetadas, localizadas entre a mácula e a periferia da olho, com rajadas de um laser de alta intensidade. A extensão do tratamento varia de acordo com a doença e com o tamanho da área afetada. Por vezes são necessárias centenas de aplicações do laser (Figura 1).



Figura 1 – Exemplos dos resultados de tratamentos por fotocoagulação a laser (AABERG, 2013; KAPLAN; PIRI, 2013).

O microscópio com lâmpada de fenda é amplamente utilizado no diagnóstico e tratamento de doenças que afetam a retina. O aparelho consiste em uma fonte de luz, capaz de iluminar o globo ocular com um feixe de luz de alta intensidade, acoplada a um microscópio. O conjunto, apresentado na Figura 2, permite ao médico examinar tanto as estruturas presentes nas camadas mais externas do globo ocular,

como a córnea e a íris, quanto a retina.

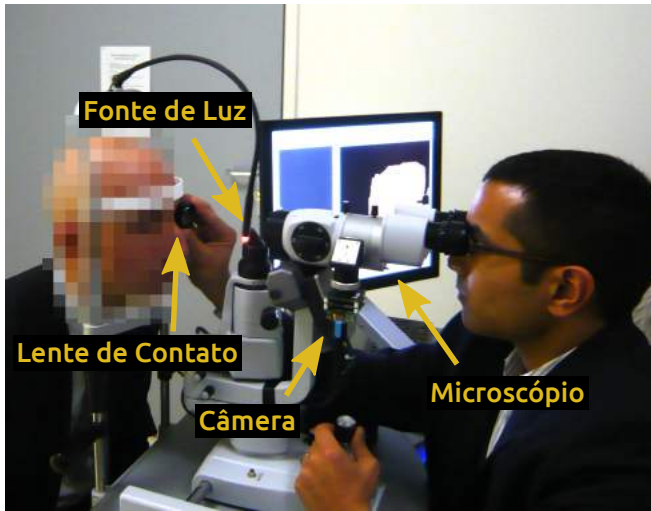


Figura 2 – Estrutura do microscópio com lâmpada de fenda. A câmera fixada ao aparelho permite o registro de todo o procedimento para análise posterior e pós-processamento.

Embora a lâmpada de fenda proporcione imagens de alta definição, seu campo de visão é limitado. Apenas uma pequena área da retina é visível ao médico. Dentre as principais dificuldades encaras durante os exames estão as frequentes perturbações causadas por reflexos da fonte de luz e movimentos do paciente. Neste contexto, a assistência computacional tem grande potencial para aumentar a segurança e eficiência do tratamento, através da transmissão de informações relevantes como testes complementares pré-operatórios, e do processamento das imagens obtidas durante exames.

Na literatura, trabalhos como o apresentado por Asmuth et al. (2001) mostram a viabilidade da construção de um mosaico a partir do registro de imagens obtidas por um microscópio com lâmpada de fenda. O resultado é um mapa que proporciona ao médico uma visão ampla de todas as estruturas da retina, mostrando a extensão total de quaisquer áreas danificadas. Desde então, diversos métodos de mapeamento de imagens médicas têm sido propostos nos contextos de cirurgias minimamente invasivas (TOTZ et al., 2011), endomicroscopia confocal (ATASOY et al., 2008) (ROSA et al., 2013), cirurgias vítreo-retinianas (RICHA et al., 2012) e fundoscopia (BROEHAN et al., 2011).

No contexto da retina, apesar do grande número de métodos de mapeamento existentes na literatura, a tendência geral é focar em imagens obtidas através de câmeras de fundo (BEHRENS et al., 2010; CAN et al., 2002a). Nesta modalidade, estruturas salientes como nervos e vasos sanguíneos são claramente visíveis, o que facilita o seu uso como referência para o registro (Figura 3). No caso da lâmpada de fenda, devido ao alto zoom da câmera, é frequente a ocorrência de imagens sem estruturas salientes, o que limita a aplicação destes métodos. Como alternativa, há trabalhos que utilizam métodos diretos, ou seja, que extraem informação diretamente dos pixels de cada imagem e a usam para o registro (ASMUTH et al., 2001; RICHIA et al., 2014). Mais detalhes sobre as vantagens e desvantagens de cada método são dados na Revisão Bibliográfica (Seção 2).

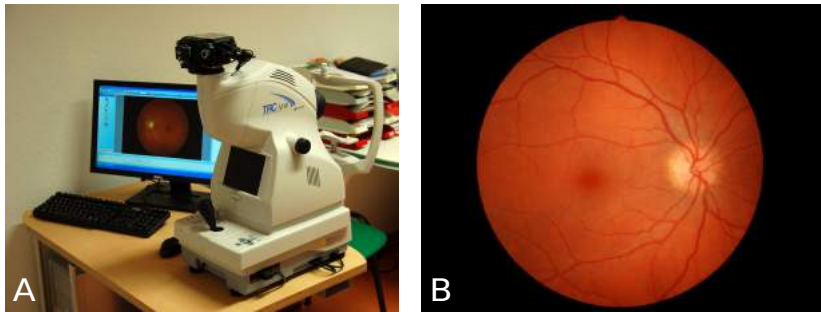


Figura 3 – A) câmera de fundo. B) imagem da retina produzida por uma câmera de fundo. Nota-se que a árvore de vasos sanguíneos é facilmente identificável.

1.1.1 Desafios

A construção de um mosaico a partir de imagens individuais envolve a superação de diversos desafios inerentes ao uso da lâmpada de fenda. A alta intensidade da fonte de luz resulta em um grande número de reflexos, por vezes obstruindo completamente o campo de visão. Bolhas formadas pelo gel aplicado à lente de contato também limitam a visibilidade. Situações como as retratadas na Figura 4 são comuns e dificultam o registro, tornando uma etapa de pré-processamento necessária para impedir que a qualidade final do mosaico seja afetada.

Além das limitações do campo de visão, alterações na intensidade

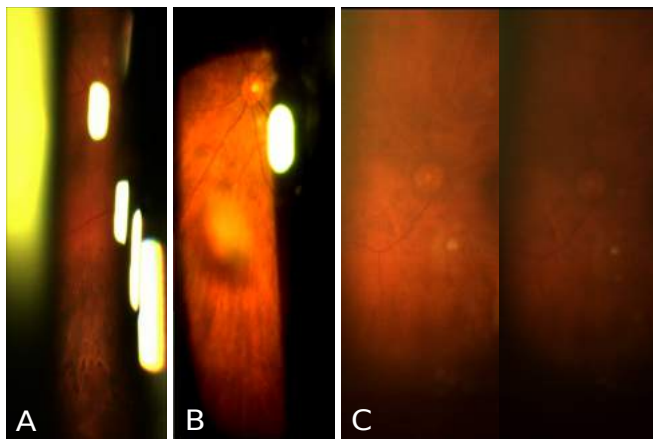


Figura 4 – Desafios comuns em imagens provenientes da lâmpada de fenda. A) Reflexos da fonte de luz. B) Obstrução causada por uma bolha do gel da lente de contato. C) Variação brusca de luminosidade.

da fonte de luz e variações temporais causam diferenças de iluminação entre as imagens. Uma mesma região da retina pode ser vista sob diversas condições de iluminação durante um único exame (Figura 4 - C), o que exige uma estratégia de registro com alta invariabilidade a condições luminosas. Diferenças de iluminação também causam bordas visíveis entre as imagens registradas, prejudicando a qualidade do mosaico final.

Durante um exame, mudanças bruscas nas condições das imagens também podem ocorrer devido a movimentos do globo ocular. Embora a cabeça do paciente seja fixada ao microscópio, o desconforto causado pela lente de contato pode causar movimentos involuntários. Durante a construção do mosaico, deve-se considerar que os atributos das imagens podem variar repentinamente.

O registro sequencial de um número elevado de imagens também sofre devido ao acúmulo de erros proveniente de imperfeições no alinhamento. O impacto desses erros torna-se evidente à medida que novas imagens são incorporadas, principalmente em exames de longa duração. O mapa resultante apresenta estruturas desalinhadas e outros erros como borrões e *ghosting* (duplicação das estruturas da imagem), causados pela junção de pares de imagens posicionadas erroneamente (Figura 5). Isso torna necessário um procedimento de pós-processamento para alinhar as imagens presentes no mosaico de modo que o resultado

final seja útil.



Figura 5 – Falha de alinhamento causado pelo acúmulo de erros no registro.

1.2 OBJETIVO GERAL

O objetivo dessa dissertação é demonstrar que um método de ajuste global é adequado para a correção de falhas como desalinhamentos, borrões e *ghosting*, presentes em mosaicos da retina construídos a partir de imagens obtidas por uma lâmpada de fenda.

1.2.1 Objetivos Específicos

As situações citadas na Seção 1.1.1 dificultam tanto o processo de construção do mosaico como o de ajuste global. Variações de luminosidade e obstruções podem prejudicar o resultado final, e devem ser tratadas durante o alinhamento. Também é necessário evitar deformar as imagens excessivamente durante o ajuste, impedindo assim que o mapa produzido seja infiel à realidade. Os objetivos específicos dessa dissertação, portanto, referem-se ao tratamento de situações adversas que potencialmente alterariam o resultado do método.

- Produzir um mosaico sem falhas resultantes do acúmulo de erros durante o mapeamento.
- Produzir um mosaico fiel à realidade, sem deformações excessivas.
- Produzir um método capaz de alinhar imagens com possíveis diferenças bruscas de iluminação.

1.3 CONTRIBUIÇÕES

A principal contribuição dessa dissertação é um método de alinhamento de mosaicos da retina construídos a partir de imagens obtidas pelo microscópio com lâmpada de fenda. As estratégias de registro propostas na literatura focam em alinhar imagens de câmeras de fundo, que permitem a compactação de informação em poucos pontos-chave, o que reduz o custo computacional. As imagens da lâmpada de fenda, no entanto, possuem um número limitado de estruturas salientes, o que as impede de serem descritas com facilidade. O método proposto é capaz de alinhar mapas com sucesso, corrigindo anomalias causadas pelo acúmulo de erros durante o mapeamento, mesmo em imagens nas quais o uso de pontos-chave é inviável.

A pesquisa realizada durante o desenvolvimento dessa dissertação

culminou em diversas publicações, a primeira sendo referente ao *framework* de mapeamento que serve de base para o processo de ajuste fino (RICHA et al., 2014). O *framework* apresenta uma abordagem ao mapeamento da retina com base em imagens da lâmpada de fundo, em contraste à grande maioria dos trabalhos na literatura, e é capaz de gerar mapas de alta qualidade, com alto nível de detalhes. O *framework* em seguida foi estendido através da adição de um método de redetecção, visando reinicializar o processo de mapeamento após interrupções causadas por obstruções (LINHARES et al., 2014). A estratégia de ajuste global em si também foi publicada e recebeu o prêmio de melhor artigo escrito por um estudante na conferência *IEEE 29th International Symposium on Computer-Based Medical Systems* (LINHARES et al., 2016).

2 REVISÃO BIBLIOGRÁFICA

A pesquisa e revisão da literatura relevante foram feitas com base no modelo de revisão sistemática proposto por Kitchenham e Charters (2007). O trabalho propõe um método organizado de pesquisa bibliográfica, que busca resultados justos e sem viés. O modelo de Kitchenham e Charters (2007) é usado para resumir a literatura existente referente a um tópico específico, assim como identificar possíveis áreas de pesquisa inexploradas.

A base para uma revisão sistemática bem sucedida é identificar a pergunta de pesquisa, o que engloba saber especificar fatores como viabilidade, risco e custo das tecnologias e métodos existentes na área de interesse. A pergunta deve ser estruturada de forma a considerar características de cada trabalho, como o contexto no qual ele pode ser aplicado, o público alvo e suas contribuições.

Os conceitos resumidos acima são intencionalmente vagos e abrangentes, pois objetiva-se o seu uso em diversas linhas de pesquisa distintas. Nessa dissertação, são buscados métodos com contribuições na área mais específica de assistência computacional na oftalmologia. Tendo isso em mente, o método de Kitchenham e Charters (2007) foi aplicado nesse contexto, e formulou-se uma série de perguntas de pesquisa que buscam textos apresentando contribuições relevantes a esse tópico:

- Quais são os trabalhos existentes na literatura que utilizam o processamento de imagens para fornecer assistência computacional à medicina oftalmológica no tratamento e diagnóstico de doenças da retina?
- Quais são as características das imagens utilizadas por estes trabalhos? Elas são semelhantes às características encontradas em imagens da retina?
- Quais são as vantagens e desvantagens dos métodos apresentados nestes trabalhos do ponto de vista de imagens oftalmológicas?

Usando essas perguntas como base para a busca, foi possível encontrar um grande número de publicações apresentando métodos de assistência computacional à medicina. A literatura apresenta uma extensa busca por métodos de mapeamento de imagens, muitas vezes em conjunto com a apresentação de informações extras com o intuito de auxiliar cirurgiões durante exames e tratamentos. As modalidades de

imagens usadas variam devido a particularidades de cada instrumento, o que influi diretamente no tipo informação extraída e analisada. Esses tópicos são explanados detalhadamente nas seções seguintes.

2.1 MAPEAMENTO DE IMAGENS MÉDICAS

O problema de mapeamento de imagens médicas é abordado com sucesso em outros contextos, como cirurgias minimamente invasivas (TOTZ et al., 2011), endomicroscopia confocal (ATASOY et al., 2008; ROSA et al., 2013), cirurgias vítreo-retinianas (RICHA et al., 2012) e fundoscopia (BROEHAN et al., 2011). Há um alto grau de variabilidade entre as modalidades de imagens médicas usadas, devido ao grande número de instrumentos de diagnóstico e tratamento. Consequentemente, as características usadas como base para o mapeamento variam de acordo com fatores como textura, luminosidade e qualidade das estruturas presentes nas imagens.

Um aspecto fundamental do mapeamento é o tipo de informação usada como base para o registro de um par de imagens. A escolha dos métodos de extração e representação dos dados influi diretamente na qualidade dos resultados e no esforço computacional necessário. Os métodos presentes na literatura seguem duas vertentes, de acordo com a modalidade das imagens usadas. Imagens repletas de estruturas salientes, como nervos, vasos sanguíneos e tumores, facilitam a extração e o casamento de pontos de interesse, permitindo representar a informação de forma compacta e diminuir o esforço computacional. Em imagens mais homogêneas, no entanto, o número de pontos salientes é muitas vezes insuficiente para se executar um registro de qualidade, logo é necessário usar toda a informação disponível para o alinhamento. Os métodos de mapeamento presentes na literatura foram, portanto, divididos de acordo ao tipo de dado usado.

2.1.1 Métodos Baseados em Pontos de Interesse

Pontos de interesse (ou pontos-chave) são descritores extraídos a partir de regiões altamente representativas de uma imagem, com o intuito de representar estruturas salientes. O uso de pontos de interesse permite representar a informação disponível de forma compacta, diminuindo o custo computacional de seu processamento. Os pontos mais representativos são destacados através da aplicação de um filtro

que busca as regiões mais salientes da imagem, comumente denominado detector.

Após a detecção, descritores são gerados para representar cada ponto de interesse de modo a mantê-los únicos e comparáveis entre si. Os descritores buscam resumir a riqueza de dados presente em cada ponto saliente de modo a facilitar o processamento, simultaneamente se mantendo únicos e resistentes a variações sofridas pela imagem, como alterações na perspectiva, luminosidade ou contraste.

Métodos baseados em pontos de interesse permitem a representação de um determinado objeto na forma de um conjunto de descritores. Um descritor pode conter dados referentes a um ponto-chave e à sua região na imagem, como orientação, intensidade e cor. Nesse contexto, o trabalho publicado por Lowe (2004) apresenta uma aplicação dos pontos de interesse na detecção de objetos conhecidos em imagens. O método, denominado *Scale-Invariant Feature Transform* (SIFT), consiste em detectar os pontos-chave e comparar seus descritores aos de um conjunto de imagens contendo os objetos desejados. Os descritores usados, robustos a rotações, oclusões, e variações de luminosidade, permite a detecção dos objetos mesmo em situações adversas (Figura 6).



Figura 6 – Exemplo de detecção de objetos através do método SIFT de detecção de pontos de interesse e construção de descritores. Esquerda: objetos de interesse. Centro: imagem contendo os objetos de interesse, alterados por transformações de rotação, escala e perspectiva. Direita: instâncias dos objetos de interesse detectadas, com os pontos de interesse em destaque.

Além do método de Lowe (2004), que detecta os pontos-chave através de filtros de diferenças de Gaussianas, a literatura contém diversas outras estratégias de detecção, como as baseadas em eigenvalores da matriz de segundo momento (HARRIS; STEPHENS, 1988) e determi-

nantes de matrizes Hessianas e Laplacianas (LINDENBERG, 1998). A criação dos descritores também varia, e os dados resumidos em um descritor influem na robustez de cada método a transformações sofridas pelas imagens analisadas. Há descritores criados com base em derivadas de Gaussianas (FLORACK et al., 1994), histogramas de gradientes locais (LOWE, 2004) e respostas de ondaletas de Haar (BAY et al., 2008).

No contexto da medicina, as vantagens dos pontos de interesse viabilizam o seu uso como base para o processamento de imagens voltado a técnicas de assistência computacional. Grande parte dos trabalhos tira proveito disso para prover orientação espacial, na forma de mapas construídos a partir de imagens obtidas durante exames, onde pontos-chave são usados como referência no registro.

O trabalho de Behrens et al. (2010) apresenta uma estratégia de assistência computacional para cirurgias minimamente invasivas da bexiga. Durante o procedimento, o tecido analisado é tingido por um corante que destaca a vasculatura ao ser iluminado por uma luz azul, o que resulta em um aumento considerável no contraste entre tumores e o tecido em volta. Esta técnica, no entanto, exige que o endoscópio seja posicionado a menos de um centímetro da parede da bexiga. O campo de visão limitado resultante dificulta a navegação e a re-deteção de tumores.

Para contornar o problema citado acima, Behrens et al. (2010) propõe uma estratégia de mapeamento capaz de registrar imagens sobrepostas, construindo um mapa de toda a área explorada durante o exame. Para isso, é usado o método SURF (*Speeded-up Robust Features*) de detecção e descrição de pontos de interesse (BAY et al., 2008), devido à sua alta discriminabilidade e baixo custo computacional. O alto contraste possibilitado pelo corante facilita a detecção de um grande número de pontos salientes nas imagens, como mostra a Figura 7, a qual se refere a imagens da bexiga.

Após a detecção dos pontos e criação dos descritores, o registro é realizado através do casamento de descritores entre pares de imagens. Devido à região de sobreposição em cada par, assume-se que descritores altamente semelhantes representam a mesma área da imagem. O resultado do casamento é demonstrado na Figura 8.

O método proposto por Behrens et al. (2010) mostra a possibilidade de se registrar uma sequência de imagens durante uma cirurgia, com o objetivo de orientar o cirurgião, compensando pelas dificuldades impostas pela prática de procedimentos minimamente invasivos. No contexto da oftalmologia, grande parte dos trabalhos de mapeamento se concentra em imagens obtidas por câmeras de fundo (CAN et al., 2002a;

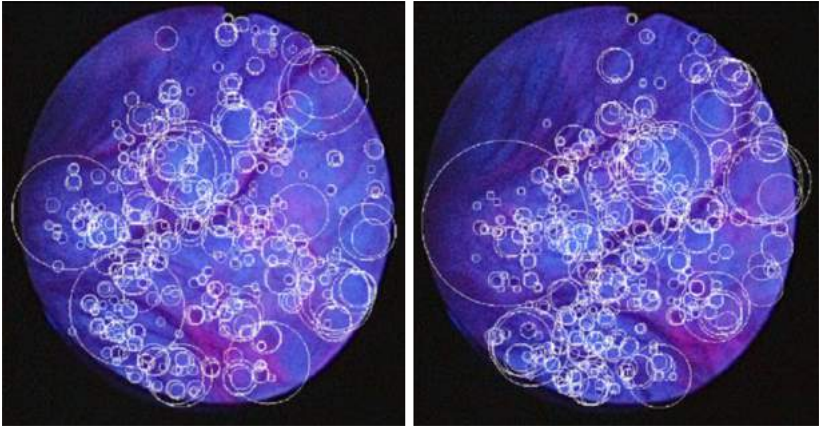


Figura 7 – Imagens obtidas durante uma cirurgia minimamente invasiva da bexiga. Em destaque, os pontos de interesse extraídos através do método proposto por Bay et al. (2008). Nota-se que, devido ao alto contraste entre as estruturas presentes no tecido, é possível detectar um grande número de pontos.

BONTALA; SIVASWAMY; PAPPURU, 2012; CHOE et al., 2006). Nesse caso, o campo de visão proporcionado é maior, o que torna possível visualizar uma área mais ampla, contendo grande parte das estruturas do olho, em especial a árvore de vasos sanguíneos. Os métodos de mapeamento de imagens de fundo tomam os vasos como ponto de referência devido ao seu alto contraste com relação ao resto da retina, assim como sua constância ao longo do exame.

O trabalho apresentado por Can et al. (2002a) se aproveita da uniformidade da estrutura de vasos sanguíneos para o registro de imagens de fundo da retina. Bifurcações e cruzamentos de vasos são usados como base para a estimação da transformação entre imagens. Para a detecção desses pontos de interesse, um procedimento recursivo de exploração, que traça os vasos a partir de pontos semente é usado (CAN et al., 1999). Como sementes, toma-se pontos no interior dos vasos, que são encontrados através de um procedimento de detecção de bordas, aplicado a uma grade de linhas e colunas uniformemente espaçadas. Locais onde há intersecções ou onde três ou mais traços convergem são considerados pontos de interesse (denominados pontos Y). Os resultados do processo de detecção são apresentados na Figura 9.

Como mencionado, imagens de câmeras de fundo proporcionam

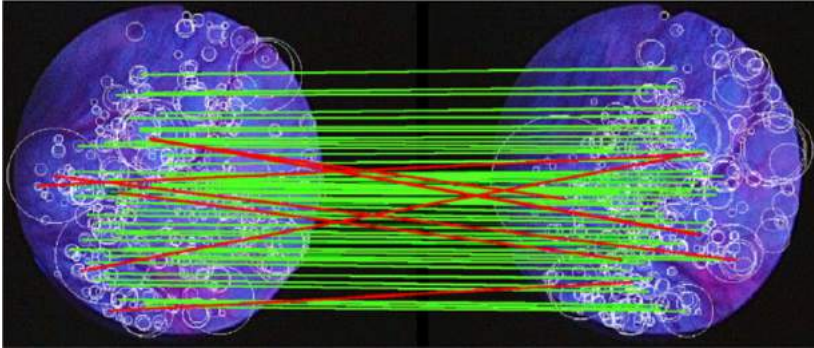


Figura 8 – Resultado do casamento de pontos de interesse entre duas imagens sequenciais da bexiga. Os pontos detectados são representados pelos círculos brancos. Linhas verdes representam pares de pontos relacionados corretamente. Linhas vermelhas são falsos positivos.

um campo de visão amplo, que cobre grande parte da retina. Isso se torna um desafio para o registro de imagens dessa modalidade devido ao formato esférico do olho. Para lidar com esse problema, o método descreve a transformação entre cada par de imagens através de um modelo de transformação quadrático. Este modelo se mostra mais adequado para representar as variações de orientação e perspectiva sofridas entre cada imagem (CAN et al., 2002b). A Figura 10 apresenta o mapa resultante do registro de uma sequência de imagens.

Bontala, Sivaswamy e Pappuru (2012) também usam pontos presentes na vasculatura para a construção de mosaicos a partir de imagens neonatais (Figura 11), obtidas durante exames de diagnóstico de retinopatia em bebês prematuros. O prognóstico depende da detecção de áreas de crescimento excessivo e fibrose na periferia da retina. Imagens neonatais frequentemente sofrem de má qualidade devido à dificuldade em sua obtenção, porém a árvore de vasos sanguíneos se mantém relativamente visível. Pontos Y, portanto, mantém sua uniformidade durante o exame.

A detecção é feita através do cálculo do determinante da Hessiana em diferentes escalas. Um limiar é aplicado aos determinantes do espaço de escala máximo para encontrar os pontos Y. O limiar escolhido é relativamente baixo para proporcionar um número viável de pontos para o registro. Um vetor de características é extraído de cada ponto com base na transformada de Radon (BATHINA; MEDATHATI; SIVASWAMY, 2010), cuja resistência a mudanças de escala, rotação e

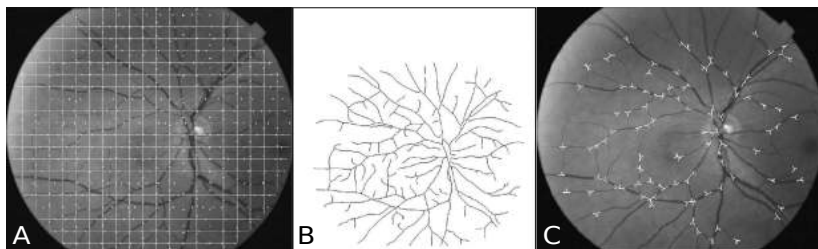


Figura 9 – Detecção de pontos de interesse pelo método de Can et al. (1999). A) grade de linhas e colunas. B) linhas traçadas pelo procedimento exploratório. C) pontos de bifurcação e cruzamento detectados.

luminosidade permite o registro de imagens altamente variáveis (Figura 11).

O conjunto de imagens de entrada também é dividido em subconjuntos, com base na posição do nervo ótico em cada imagem. Um procedimento simples de detecção do nervo ótico, com base em um modelo anelar, é usado para adicionar cada imagem a um de 5 possíveis subconjuntos, buscando maximizar a área de sobreposição de cada grupo. Após o registro dos subconjuntos, os cinco mapas resultantes são registrados, formando o mosaico final. Apesar do elevado custo computacional, resultando até 15 minutos de tempo de execução com 15 imagens 640 x 480, o método é capaz de lidar com as condições adversas impostas pelo processo de obtenção de imagens de bebês prematuros e produzir um mosaico de alta qualidade (Figura 12).

O conceito de pontos Y é estendido por Choe et al. (2006), usando toda a região onde 3 vasos convergem como base para o registro de imagens de fundo da retina. Duas modalidades de imagem são consideradas: imagens padrão da câmera de fundo e imagens de angiografia, onde os vasos sanguíneos são destacados por um corante (Figura 13). Apesar da alta variação intermodal de intensidade, a estrutura de vasos é sempre visível, portanto locais de bifurcação se mostram ideais para o alinhamento das imagens. O método usa pontos Y como sementes para um procedimento de descida de gradiente, que cria uma representação de área de convergência denominada modelo articulado.

O pressuposto para a detecção das sementes é que bordas na imagem indicam a presença de vasos sanguíneos. Candidatos a pontos Y são inicialmente obtidos através da aplicação de filtros Laplacianos de Gaussianas em 6 direções diferentes em cada *pixel*. Pontos que respondem fortemente a pelo menos 3 dos 6 filtros são considerados candidatos

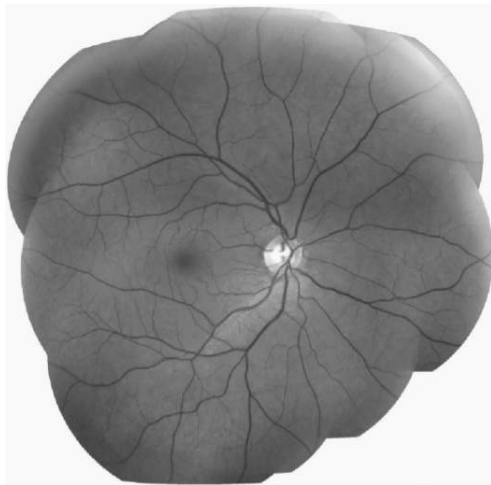


Figura 10 – Resultado do método de mapeamento de imagens de câmeras de fundo proposto por Can et al. (2002a).

a sementes e filtrados através de um procedimento de Análise de Componentes Principais. O processo de ajuste de um modelo articulado a partir de uma semente é demonstrado na Figura 14.

Nota-se que os métodos de mapeamento baseados em pontos de interesse possuem características e restrições em comum. O uso desses pontos-chave permite a compactação dos dados presentes em grandes regiões de uma imagem, resultando em uma redução no esforço computacional do seu processamento. Isso, porém, depende da qualidade dos procedimentos de detecção e descrição. Além disso, é necessário um grande número de pontos de interesse, com descritores únicos, para possibilitar um registro correto. Isso só se torna possível quando as imagens possuem estruturas salientes, como vasos sanguíneos. Em modalidades nas quais o campo de visão é limitado, como em imagens da lâmpada de fenda, fatores como obstruções ou o uso de imagens de regiões periféricas da retina podem limitar a presença dessas estruturas, impossibilitando o registro (Figura 18).

O método de mapeamento proposto por Chanwimaluang, Fan e Fransen (2006) busca lidar com as desvantagens inerentes ao uso de pontos de interesse. As imagens usadas são obtidas com base no protocolo para tratamento adiantado de retinopatia diabética, proposto pelo Instituto Nacional de Saúde (NHI) americano, para tratamento adian-

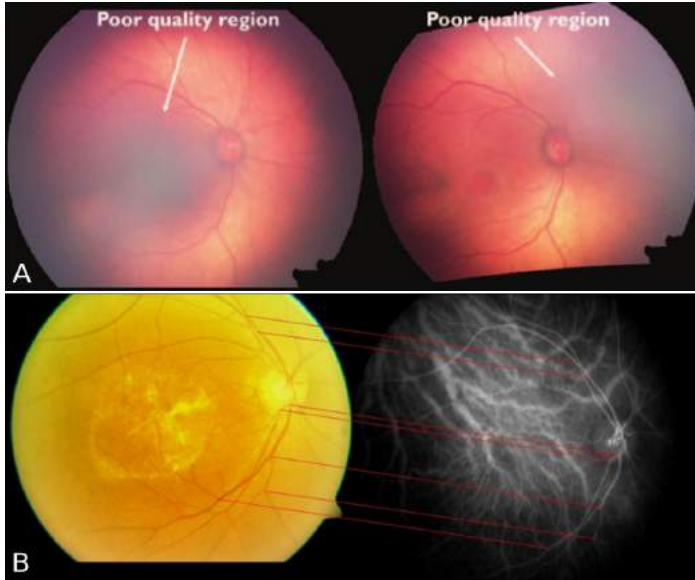


Figura 11 – A) Exemplos de imagens neonatais da retina, obtidas por uma câmera de fundo. Em destaque, as áreas de baixa qualidade que obstruem o campo de visão, perturbando o registro. B) Correspondência entre pontos Y entre uma imagem de fundo (esquerda) e uma imagem angiográfica (direita), usando os descritores apresentados por Bathina, Medathathi e Sivaswamy (2010).

tado de retinopatia diabética (National Institute of Health (NIH), 2014), que define 7 áreas de uma imagem de fundo cobrindo campos específicos da retina (Figura 15). A estrutura de vasos sanguíneos é visível em todas as áreas, mas métodos baseados em pontos de interesse sofrem com as distribuições não uniformes de intensidade, resultantes da aquisição imperfeita das imagens. Nota-se também que algumas das regiões possuem grandes porções homogêneas, sem vasculatura ou textura, que limitam o número de pontos disponíveis.

Com o objetivo de lidar com imagens com ausência de pontos de interesse, a estratégia proposta incorpora o uso de métodos diretos, ou seja, métodos em que se extrai parâmetros diretamente de cada pixel da imagem. As duas estratégias são aplicadas hierarquicamente, iniciando com uma transformação simples, cuja complexidade aumenta à medida que os parâmetros são estimados. A árvore de vasos sanguíneos é a



Figura 12 – Resultado do método de mapeamento proposto por Bontala, Sivaswamy e Pappuru (2012), aplicado a de câmara de fundo obtidas de bebês prematuros. Mapa formado a partir de um conjunto de 31 imagens, das quais 16 foram descartadas devido a má qualidade.

base para o processo hierárquico, portanto primeiramente ela é extraída através de um método de limiarização, com base em uma medida de entropia local da imagem (CHANWIMALUANG; FAN, 2003b, 2003a). O resultado é uma máscara binária correspondente aos vasos (Figura 16 - B).

Para a extração dos pontos Y, a árvore binária é erodida, resultando em um esqueleto formado apenas pelas linhas centrais dos vasos (Figura 16 - C). Em seguida, candidatos a pontos Y são detectados por um *kernel* 3x3, que marca as posições cuja vizinhança possui pelo menos 3 pixels pertencentes a uma linha central. Por último, um *kernel* 11x11 é aplicado para descartar candidatos inadequados. Se um candidato possui pelo menos 2 pontos pertencentes a uma linha central na borda do *kernel*, ele é marcado como ponto Y. O resultado da detecção é mostrado na Figura 16 (D).

Um método direto baseado na minimização de uma medida de informação mútua entre as imagens é usado para estimar a transformação de translação aplicada aos pontos de interesse. Em seguida, as trans-

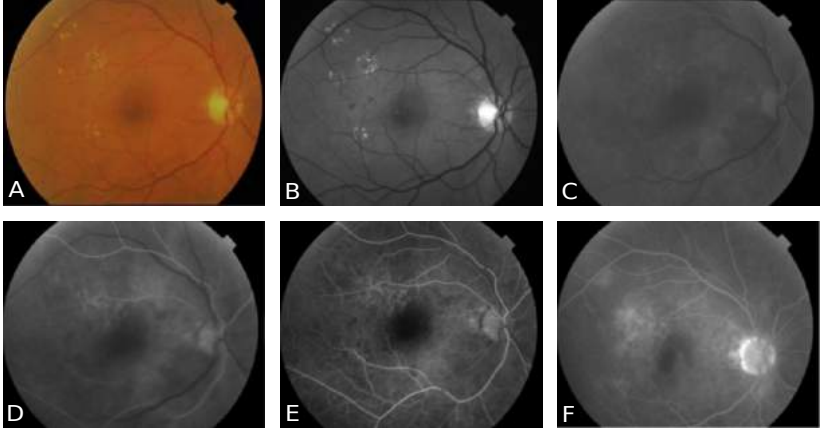


Figura 13 – A e B mostram imagens padrão obtidas pela câmera de fundo. C, D, E e F mostram o efeito da aplicação de fluoresceína. Nota-se a alteração na coloração dos vasos à medida que o corante circula pelo sangue.

formações de escala e rotação são estimadas a partir dessas coordenadas atualizadas, minimizando-se uma medida de erro:

$$M = \arg \min_M \min_{q \in Q} \|p - T(q; M)\|^2 \quad (2.1)$$

onde p e q denotam os conjuntos de pontos de interesse de duas imagens, e $T(q; M)$ representa as coordenadas do ponto q transformadas pelo modelo M . O procedimento iterativo é refinado através do descarte de pares de pontos cuja distância Euclidiana se torna grande demais. O resultado do processo hierárquico de registro é mostrado na Figura 17.

Apesar do uso de um método direto em conjunto com os pontos de interesse, a estratégia de Chanwimaluang, Fan e Fransen (2006) ainda depende da detecção de um conjunto de pontos-chave para o registro. O uso deste algoritmo, portanto, ainda se torna inviável caso a detecção falhe em encontrar um número suficiente de pontos salientes. Devido ao campo de visão pouco extenso de ferramentas como a lâmpada de fenda, o número de pontos-chave é limitado em regiões sem a presença de estruturas salientes (Figura 18 - B). Nesses casos, utilizar pontos salientes como referência para o alinhamento tem resultados pouco confiáveis e o uso de toda a informação disponível na imagem é necessário. Isso é abordado com mais detalhes na seção seguinte.

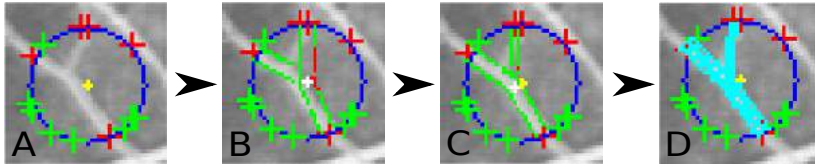


Figura 14 – O procedimento de ajuste de um modelo articulado por descida de gradiente. A mostra a posição do ponto Y. As iterações gradualmente deformam o modelo, como mostrado em B e C. O modelo ajustado é mostrado em D.

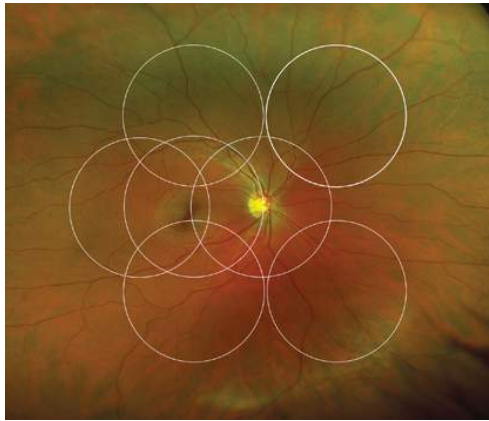


Figura 15 – Sete campos da retina definidos pelo Instituto Nacional de Saúde (NHI) americano, no estudo sobre tratamento adiantado de retinopatia diabética (National Institute of Health (NIH), 2014).

2.1.2 Métodos Diretos

Em contraste aos métodos baseados em pontos de interesse, métodos diretos extraem parâmetros diretamente de valores mensuráveis em cada pixel ou região de uma imagem, buscando utilizar toda a informação disponível. Esse tipo de estratégia descarta a compactação da informação oferecida pelos pontos de interesse, resultando em uma desvantagem inerente de ter que processar mais dados. No entanto, como explicado na Seção 2.1.1, é possível que não haja pontos de interesse suficientemente únicos para um registro confiável, principalmente em imagens médicas. O alinhamento de imagens através de métodos

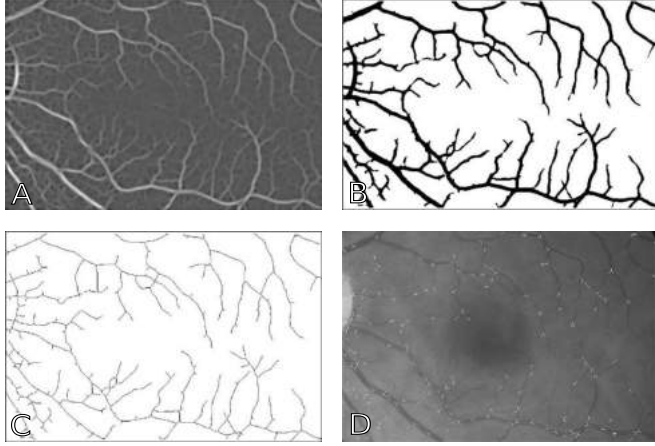


Figura 16 – Extração da árvore binária e detecção de pontos Y (CHANWIMALUANG; FAN; FRANSEN, 2006). A) imagem de entrada. B) árvore binária. C) linhas centrais da árvore. D) pontos Y detectados.

diretos não possui essa restrição.

O trabalho clássico de Lucas e Kanade (1981), expandido por Baker e Matthews (2001), modela o registro de duas imagens como um problema de busca por subida de encosta. Dadas duas imagens I e T , o objetivo do algoritmo é minimizar a seguinte função de similaridade:

$$\sum_{\mathbf{x}} (I(w(\mathbf{x}, \mathbf{p})) - T(\mathbf{x}))^2 \quad (2.2)$$

onde $w(\mathbf{x}, \mathbf{p})$ representa uma função que transforma cada pixel \mathbf{x} em I , segundo um modelo de transformação com um vetor de parâmetros \mathbf{p} . O método assume que há uma estimativa inicial dos parâmetros de transformação, e a minimização da Equação 2.2 é feita iterativamente, com respeito a \mathbf{p} :

$$\sum_{\mathbf{x}} (I(w(\mathbf{x}, \mathbf{p} + \Delta\mathbf{p})) - T(\mathbf{x}))^2 \quad (2.3)$$

$$\mathbf{p} = \mathbf{p} + \Delta\mathbf{p} \quad (2.4)$$

onde $\Delta\mathbf{p}$ é a variação estimada dos valores em \mathbf{p} em cada iteração. As equações acima são executadas até que o vetor convirja, ou seja, até que $\Delta\mathbf{p}$ seja inferior a um limiar ϵ .

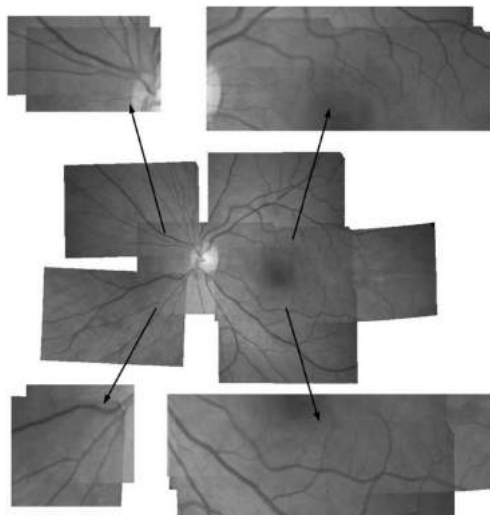


Figura 17 – Resultados da estimação da transformação projetiva entre os 7 campos definidos pelo protocolo ETDRS, com algumas áreas de sobreposição ampliadas.

Desde sua publicação, aplicações do método original de Lucas e Kanade (1981) foram propostas em áreas como rastreamento (BLACK; JEPSON, 1998; HAGER; ; BELHUMEUR, 1998), estimação de movimento (BERGEN et al., 1992) e codificação de faces (COOTES; EDWARD; TAYLOR, 1998). No contexto de imagens da retina, no entanto, o número de aplicações na literatura é limitado. Devido ao amplo campo de visão oferecido pelas câmeras de fundo, o alinhamento de imagens com base em pontos de interesse localizados na árvore de vasos sanguíneos se mostra mais viável. O registro de imagens da lâmpada de fenda, entretanto, resulta em mapas mais detalhados, devido à alta definição das imagens desta modalidade. Em procedimentos que exigem precisão, como a fotocoagulação a laser, isso é vantajoso por ter um potencial de assistência computacional maior para o cirurgião.

O uso de métodos diretos em imagens da lâmpada de fenda foi introduzido por Asmuth et al. (2001). O método fixa uma imagem inicial e a incrementa com as imagens seguintes (Figura 19 - esquerda), realizando o registro através da minimização por Gauss-Newton, segundo descrito por Lucas e Kanade (1981), formando um mosaico completo da retina. Para remover as bordas resultantes de variações de iluminação,

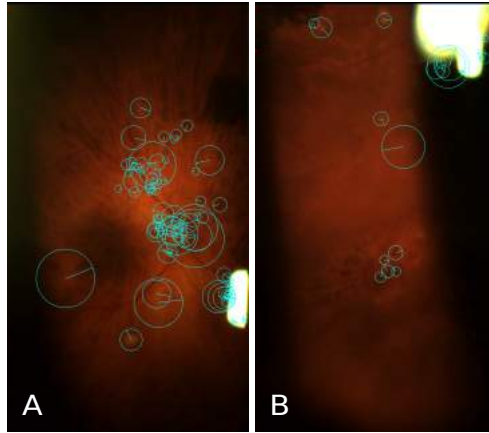


Figura 18 – Pontos SURF detectados em imagens da lâmpada de fenda. A) imagem com alto contraste entre a árvore de vasos sanguíneos e a retina. B) imagem de uma região com pouca informação de textura e sem estruturas salientes. Nota-se o baixo número de pontos presentes.

cada imagem incorporada é decomposta em uma representação por pirâmides Laplacianas. O mosaico final, com as imperfeições removidas, é o resultado da reconstrução das imagens a partir das pirâmides Laplacianas (Figura 19 - direita) (BURT; ADELSON, 1983).

No mesmo contexto, um *framework* para a construção de um mosaico durante exames com lâmpadas de fenda foi proposto em colaboração com o autor dessa dissertação (RICHA et al., 2014). O mapeamento consiste em sobrepor as imagens individuais em uma grade, que é expandida através da adição de novas imagens à medida que a retina é explorada. Durante o procedimento, uma das imagens na grade é rastreada através da estratégia de minimização de uma função de similaridade apresentada por Lucas e Kanade (1981). A Figura 20 (A) mostra os resultados da aplicação do *framework*.

O resultado da técnica de Richa et al. (2014) mostra o grande potencial do uso de métodos diretos para o rastreamento e mapeamento de imagens da lâmpada de fenda. Entretanto, o acúmulo de erros de rastreamento é uma falha inerente a métodos de aproximação, como o proposto por Lucas e Kanade (1981). Em exames extensos o mapa passa a mostrar desalinhamentos (Figura 20 - B). Isso influi na capacidade do médico de realizar procedimentos cirúrgicos e exames pós-operatórios a partir de informações confiáveis.

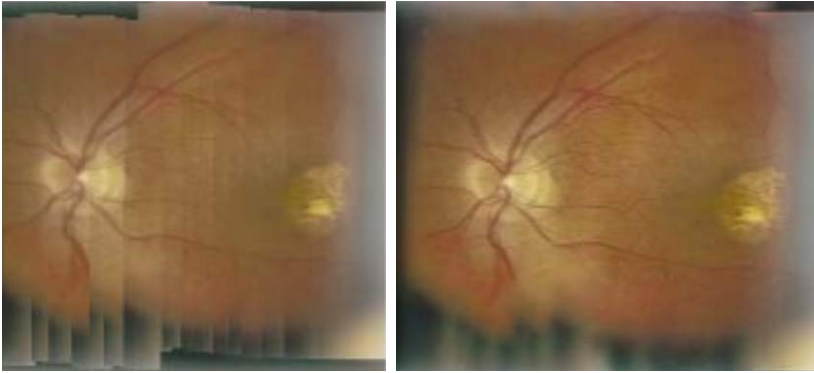


Figura 19 – Resultado da técnica de mapeamento da retina proposta por Asmuth et al. (2001). Esquerda: mosaico inicial, contendo bordas evidentes causadas por variações de iluminação entre imagens. Direita: mosaico final, resultante da junção das imagens a partir da reconstrução das pirâmides Laplacianas.

Além do acúmulo de erros, o rastreamento também pode ser interrompido por obstruções causadas por bolhas de gel, reflexos da luz da lâmpada e movimentos bruscos do paciente. Isso torna necessária uma estratégia de relocalização, com o intuito de determinar a posição atual do microscópio na retina após o rastreamento ser suspenso. Com esse objetivo, o trabalho de Richa et al. (2014) é estendido por Linhares et al. (2014), no qual é apresentada uma estratégia de relocalização.

O problema de relocalização abordado consiste em estimar uma transformação indicando a posição atual da câmera no mapa da retina, permitindo assim que o rastreamento seja reiniciado. Para isso, as imagens presentes na grade que forma o mosaico e no campo de visão são representadas por padrões binários locais (LBP): descritores compactos de textura que codificam as diferenças de intensidade entre pixels vizinhos. Devido ao grande número de imagens possivelmente contidas na grade, uma representação LBP compacta de 2 bits é adotada (KALAL; MATAS; MIKOLAJCZYK, 2010; YE et al., 2013). Uma região de interesse com as dimensões das imagens na grade é extraída e um código é atribuído a cada posição x , de acordo com a posição do pixel de maior intensidade entre x e os 3 pixels adjacentes abaixo e à direita. O processo de extração dos LBP é demonstrado na Figura 21.

A similaridade entre um par de imagens I e T é representada pelo número de pixels com o mesmo código LBP:

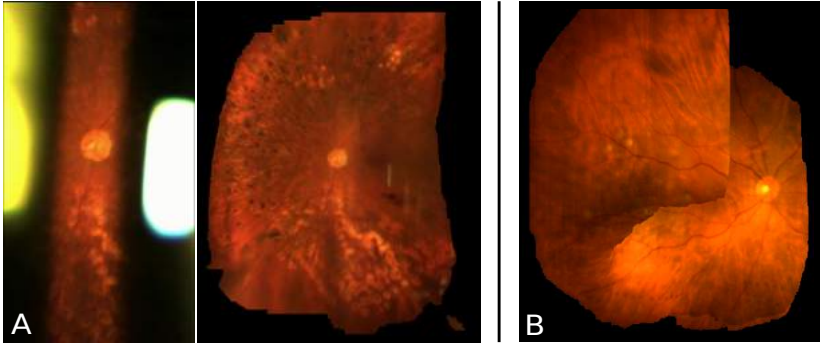


Figura 20 – A) Campo de visão da lâmpada de fenda e mosaico resultante do *framework* de mapeamento proposto por Richa et al. (2014). Nota-se a alta qualidade dos mosaicos gerados, apesar do campo de visão estreito proporcionado pela lâmpada de fenda. B) Resultado de-salinhado, devido ao acúmulo de erros de rastreamento.

$$\mathcal{L}(I, T) = \sum_{\mathbf{x}} \left[\mathcal{C}(I(\mathbf{x})) \wedge \mathcal{C}(T(\mathbf{x})) \right] \quad (2.5)$$

onde $\mathcal{C}(I(\mathbf{x}))$ e $\mathcal{C}(T(\mathbf{x}))$ representam os códigos referentes aos pixels nas posições \mathbf{x} das imagens I e T , e \wedge representa uma comparação, isto é, $\mathcal{C}(I(\mathbf{x})) \wedge \mathcal{C}(T(\mathbf{x})) = 1$ se $I(\mathbf{x})$ tem o mesmo código que $T(\mathbf{x})$, e 0 caso contrário.

A realocização é realizada através da comparação par-a-par da região de interesse a cada imagem no mapa, com base na medida de similaridade (Equação 2.5). A imagem da grade mais similar à região de interesse é usada como referência para a reinicialização do rastreamento. O uso dos códigos baseados em padrões binários locais de 2 bits possibilita uma comparação com um esforço computacional mínimo, muito inferior ao de métricas de similaridade mais complexas como a correlação cruzada (LOVEGROVE, 2011) e informação mútua (DAME; MARCHAND, 2012).

A pesquisa sistemática da literatura mostra uma grande tendência ao processamento de imagens obtidas por câmeras de fundo. Já que elas possuem a vantagem de mostrar uma grande área da retina, o foco de trabalhos buscando oferecer alternativas de assistência computacional é aproveitar este campo de visão amplo e usar pontos Y (CAN et al., 2002a; BONTALA; SIVASWAMY; PAPPURU, 2012; CHOE et al., 2006) como base para a criação de um mapa. Em procedimentos nos quais a mani-

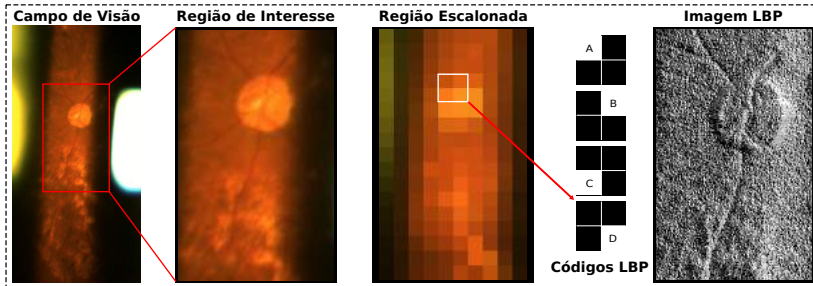


Figura 21 – Exemplo da extração dos códigos LBP de uma imagem. Uma região de interesse é gerada a partir do campo de visão atual da câmera. Cada pixel da região de interesse é comparada aos pixels adjacentes e um código que representa a posição do pixel mais intenso é atribuído.

pulação do tecido em questão é viável, também é possível destacar as estruturas salientes através de corantes (BEHRENS et al., 2010). Ambas essas vantagens são ausentes em exames oftalmológicos com câmera de fenda, no entanto, e portanto poucos são os trabalhos propondo o mapeamento da lâmpada de fenda. Mesmo os métodos existentes sofrem de falhas decorrentes da impossibilidade do uso desses pontos de interesse (RICHA et al., 2014). Por esses motivos, essa dissertação apresenta um método que aumenta o potencial de assistência computacional em procedimentos médicos feitos com lâmpadas de fenda através do ajuste dos mapas resultantes.

O trabalho proposto por Richa et al. (2014) foi escolhido como base para o desenvolvimento desse método, pois organiza as imagens obtidas em uma grade facilmente acessável. Apesar de o *framework* base ser capaz de contruir o mosaico simultaneamente ao próprio exame sendo realizado, a adição de um procedimento de ajuste global baseado em métodos diretos aumenta o esforço computacional necessário, tornando a execução simultânea proibitiva. Além disso, a constante adição de novas imagens ao mapa durante com a progressão do exame torna o processo imprevisível. Por esses motivos, o custo computacional relativo ao ajuste global está fora do escopo dessa dissertação. Na Seção 3.1 são dados detalhes sobre a implementação de ambos o *framework* de Richa et al. (2014) e do ajuste global.

3 METODOLOGIA

Como citado, métodos de mapeamento da retina sofrem deterioração devido ao acúmulo de erros de registro. Imperfeições na estimação da transformação entre a imagem de referência e a imagem variável podem resultar em erros de alinhamento, estruturas duplicadas (*ghosting*) e borrões (Figura 22). Esse trabalho busca, portanto, corrigir mosaicos da retina através de ajuste global, de modo a mitigar essas deformações. A maioria dos métodos na literatura lidam com imagens da lâmpada de fundo, que apresentam um campo de visão maior e a maior parte da árvore de vasos sanguíneos claramente visível, o que favorece o uso de métodos baseados em pontos de interesse. Imagens obtidas pela lâmpada de fenda, no entanto, frequentemente não possuem um número suficiente de pontos de interesse, portanto a aplicação de um método direto é a abordagem mais adequada.

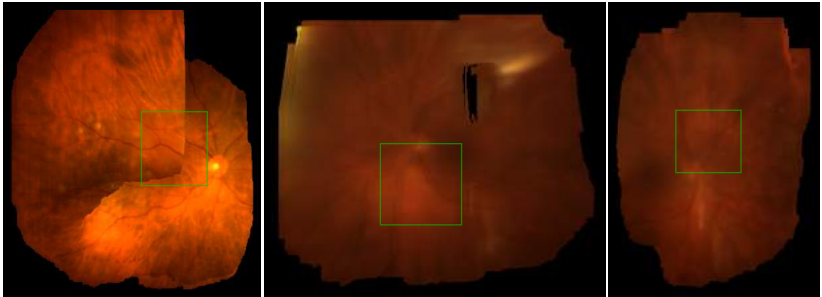


Figura 22 – Problemas causados pelo acúmulo de erros de rastreamento durante o mapeamento da retina. Esquerda: desalinhamento das estruturas da retina. Meio: duplicação de vasos sanguíneos (*ghosting*). Direita: borrões e baixa nitidez. Em destaque, regiões de particular interesse.

A contribuição principal dessa dissertação é um algoritmo de ajuste global que pode ser aplicado a imagens resultantes de um procedimento de mapeamento da retina. O objetivo do método proposto é estender a já existente estratégia de mapeamento proposta por Richa et al. (2014) de modo a consertar as falhas ocasionadas pelo acúmulo de erros inerente ao uso de métodos diretos. O algoritmo assume a existência de um conjunto de imagens resultante de um procedimento de mapeamento da retina. O deslocamento entre as imagens pode ser

fixo ou variável, porém é necessário que haja uma região considerável (pelo menos 30% da área de uma imagem) de sobreposição entre imagens vizinhas para que os resultados do ajuste sejam satisfatórios. O método usado para o mapeamento foi de Richa et al. (2014) devido à alta qualidade dos mosaicos construídos e o fato de as imagens serem organizadas em uma grade, com grandes áreas em comum entre imagens próximas. O procedimento de mapeamento é explicado na Seção 3.1. Detalhes sobre a contribuição dessa dissertação, ou seja, sobre o ajuste global, são dados na Seção 3.2.

O fluxograma na Figura 23 apresenta uma visão geral simplificada do processo, que se inicia com um exame da retina, durante o qual é realizado o mapeamento, produzindo uma grade de imagens sobrepostas que representam o mosaico. No início do exame, a região central do campo de visão atual da lâmpada de fenda é extraída e adicionada à grade como uma única imagem inicial. O rastreamento da imagem inicial é iniciado em seguida e novas imagens são adicionadas quando detecta-se um afastamento do campo de visão atual ao centro da imagem rastreada. Caso o rastreamento seja interrompido por obstruções, um procedimento de redetecção é acionado. Após a finalização do exame, a grade de imagens resultante é ajustada até que ocorra a convergência global, e o resultado final é o mosaico ajustado.

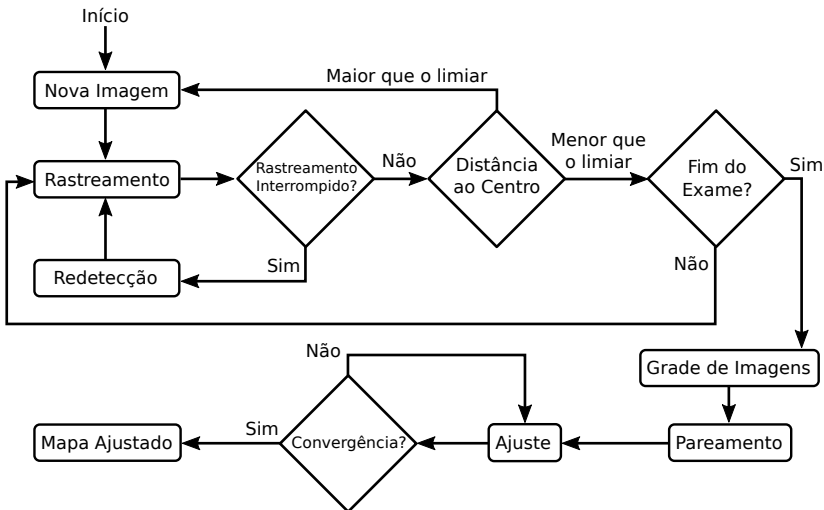


Figura 23 – Fluxograma demonstrando o funcionamento de ambos os métodos de mapeamento da retina e de ajuste global.

3.1 MAPEAMENTO DA RETINA

Uma estratégia de mapeamento baseada em métodos diretos para o registro de imagens obtidas durante um exame com a lâmpada de fenda é proposta (RICHA et al., 2014). O procedimento consiste em tomar uma imagem inicial como âncora e registrar novas imagens a ela, incrementando o mapa à medida que a retina é explorada. O resultado do mapeamento é um mosaico centrado na âncora, composto por um conjunto de imagens adjacentes (Figura 24).

O método possui uma etapa de pré-processamento, com o objetivo de lidar com as frequentes obstruções e extrair a área útil da imagem. Durante o exame, uma das imagens que compõe o mosaico é rastreada e sua posição no campo de visão é monitorada. O mosaico é estendido quando o centro da imagem rastreada no campo de visão se afasta do seu centro real no mapa da retina. O rastreamento em si é realizado através da minimização de uma função de similaridade entre uma imagem de referência e o atual campo de visão da câmera. O *framework* também usa um procedimento de seleção eficiente de pixels para reduzir o custo computacional. Detalhes sobre cada etapa do método são dados nas seções seguintes.

3.1.1 Pré-processamento

A etapa de pré-processamento é realizada com o intuito de filtrar obstruções, assim como áreas não iluminadas da retina, que podem afetar negativamente o registro. Inicialmente, uma operação de limiarização, cujo limiar é computado com base nas intensidades dos canais verde e vermelho das imagens, é aplicada. Devido à alta taxa de absorção da luz azul pelo globo ocular (Figura 25 - esquerda), o canal azul possui baixa representatividade e é desconsiderado. Um pixel é considerado parte da retina se $r - 0.7g > 0$, onde r e g são as intensidades dos seus canais vermelho e verde, respectivamente. O resultado da limiarização é refinado através da aplicação de uma operação de fechamento morfológico (dilatação seguida de erosão, com *kernels* circulares com dimensões 8x8 e 30x30, respectivamente). O resultado final da etapa de pré-processamento é apresentado na Figura 25 (direita).

Mesmo após a extração de uma região de interesse adequada, a alta taxa de variação de luminosidade inerente às imagens da lâmpada de fenda torna o registro inviável sem uma forma de compensação de iluminação. Richa et al. (2014) faz uso do método de compensação não

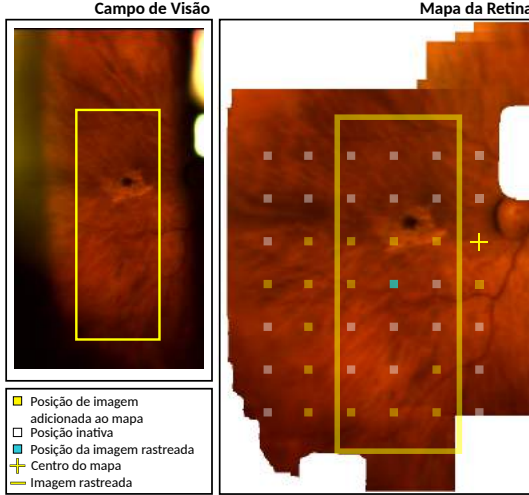


Figura 24 – Estrutura do mosaico construído através da estratégia de mapeamento proposta por Richa et al. (2014). Esquerda: campo de visão atual do microscópio. Em destaque, a imagem sendo rastreada. Direita: representação da grade de imagens que formam o mosaico, indicando a posição relativa da imagem rastreada.

rígido que modela variações locais no contraste usando uma superfície paramétrica deformável $g(x, \mu)$ (SILVEIRA; MALIS, 2010). Cada imagem I^* é compensada por uma aproximação a uma superfície *Thin-Plate Spline* (LIM; YANG, 2005):

$$I(\mathbf{x}, \mu, \beta) = g(\mathbf{x}, \mu)I^*(\mathbf{x}) + \beta \quad (3.1)$$

onde $g(\cdot)$ é uma função das posições dos pixels \mathbf{x} e o vetor de pontos de controle μ , e β é uma variável para estimar variações globais de iluminação.

3.1.2 Rastreamento

Durante o exame, a imagem mais próxima ao campo de visão atual é rastreada através de um método direto baseado na minimização de uma função de erro entre uma imagem de referência T e uma imagem variável I (LUCAS; KANADE, 1981). Dadas as posições dos pixels \mathbf{x} em T e $w(\mathbf{x}, \mathbf{p})$ uma função de transformação com parâmetros \mathbf{p} que mapeia

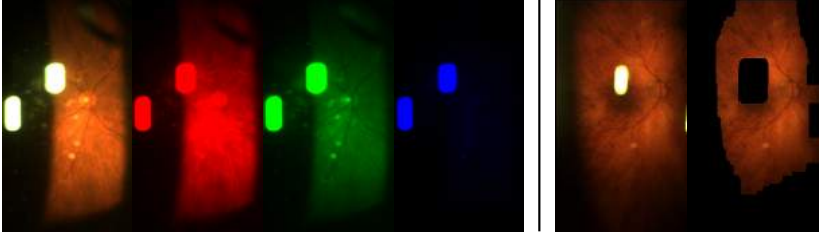


Figura 25 – Esquerda: Separação de uma imagem da lâmpada de fenda em seus canais vermelho, verde e azul. Nota-se a baixa representatividade do canal azul, devido à alta taxa de absorção da luz azul pela retina. Direita: resultado da etapa de pré-processamento.

todo pixel \mathbf{x} de T para I , o rastreamento consiste em estimar \mathbf{p} de modo a minimizar a seguinte função:

$$\min_{\mathbf{p}, \mu, \beta} \sum_{\mathbf{c}} \sum_{\mathbf{x}} \left({}^c I(w(\mathbf{x}, \mathbf{p}), \mu, \beta) - {}^c T(\mathbf{x}) \right)^2 \quad (3.2)$$

onde c indica um canal de cor. Novamente, devido à baixa representatividade do canal azul, são usados apenas os canais vermelho e verde. μ e β são os parâmetros do procedimento de compensação de iluminação.

A escolha do método de parametrização da função $w(\cdot)$, representado pelo vetor \mathbf{p} , impacta o número de graus de liberdade do alinhamento e o esforço computacional para calculá-lo. No caso da lâmpada de fenda, o paciente tem sua cabeça fixada ao suporte do microscópio, portanto mudanças na escala não são significativas. O cirurgião é livre para transladar a câmera horizontalmente e verticalmente e movimentos involuntários da cabeça do paciente causam pequenas variações de rotação, portanto o vetor de parâmetros \mathbf{p} escolhido segue um modelo de rotação e translação, capaz de lidar com todas as variações possíveis entre as imagens de entrada.

O rastreamento é feito através de um procedimento de descida de encosta, que consiste em calcular iterativamente, a partir de uma estimativa inicial dos parâmetros \mathbf{p}_0 os elementos de \mathbf{p} para os quais o resultado da Equação 3.2 é mínimo. A cada iteração, um vetor de variações $\Delta\mathbf{p}$ é estimado e \mathbf{p} é atualizado até que $\|\Delta\mathbf{p}\| < \epsilon$, onde ϵ indica um limiar mínimo de variação aceitável. Isso é feito de maneira eficiente por um método de minimização de segunda ordem (BENHIMANE; MALIS, 2007; RICHA; SZNITMAN; HAGER, 2012). Primeiramente, uma expansão de Taylor é aplicada a $I(w(\mathbf{x}, \mathbf{p}))$ com relação ao vetor de

parâmetros atual \mathbf{p}_a :

$$T(x) = I(w(\mathbf{x}, \mathbf{p}_a)) + J(\mathbf{p})\Delta\mathbf{p} + \frac{1}{2}H(\mathbf{p}, \Delta\mathbf{p})\Delta\mathbf{p} \quad (3.3)$$

onde $J(\mathbf{p})$ representa uma matriz Jacobiana:

$$J(\mathbf{p}) = \frac{\partial I(w(\mathbf{x}, \mathbf{p}))}{\partial \mathbf{p}} \quad (3.4)$$

e $H(\mathbf{p}, \Delta\mathbf{p})$ é uma matriz Hessiana que depende de $\Delta\mathbf{p}$ e da derivada de segunda ordem de $I(w(\mathbf{x}, \mathbf{p}))$ com relação a \mathbf{p} , avaliada em \mathbf{p}_a . Segundo o método de Gauss-Newton, o termo à direita na Equação 3.3 pode ser descartado para gerar uma aproximação ao valor de $\Delta\mathbf{p}$:

$$\Delta\mathbf{p} = -J(\mathbf{p}_c)^+(I(w(\mathbf{x}, \mathbf{p})) - T(x)) \quad (3.5)$$

onde $(.)^+$ representa a pseudo-inversa de uma matriz. Com o intuito de aumentar a taxa de convergência do processo iterativo, uma aproximação de segunda ordem da Hessiana pode ser obtida através de uma expansão de primeira ordem de $J(\mathbf{p})$ com relação a \mathbf{p}_a , avaliada em \mathbf{p}_0 (BENHIMANE; MALIS, 2007):

$$J(\mathbf{p}_0) \approx J(\mathbf{p}_a) + H(\mathbf{p}, \Delta\mathbf{p}) \quad (3.6)$$

onde $J(\mathbf{p}_0)$ pode ser calculado como:

$$J(\mathbf{p}_0) = \frac{\partial T(\mathbf{x})}{\partial \mathbf{p}} \quad (3.7)$$

onde \mathbf{p}_0 representa os parâmetros da transformação identidade, ou seja, os parâmetros segundo os quais $w(\mathbf{x}, \mathbf{p})$ não afeta a imagem I . Isolando $H(\mathbf{p}, \Delta\mathbf{p})$ na Equação 3.6, substituindo o resultado na Equação 3.3 e isolando $\Delta\mathbf{p}$, temos a equação que representa a variação dos parâmetros de \mathbf{p} em uma iteração do método eficiente de minimização de segunda ordem proposto por Benhimane e Malis (2007):

$$\Delta\mathbf{p} = -2(J(\mathbf{p}_0) + J(\mathbf{p}_a))^+(I(w(\mathbf{x}, \mathbf{p})) - T(\mathbf{x})) \quad (3.8)$$

Durante o mapeamento, a correlação cruzada normalizada (NCC) entre a imagem de referência T e a imagem I é usada como medida de confiança do rastreamento:

$$NCC(T, I) = \frac{\sum_x (T(x) - \bar{T})(I(x) - \bar{I})}{\sqrt{\sum_x (T(x) - \bar{T})^2} \sqrt{\sum_x (I(x) - \bar{I})^2}} \quad (3.9)$$

onde \bar{T} e \bar{I} representam a média da intensidade dos pixels nas imagens T e I , respectivamente. A correlação cruzada permite medir a similaridade entre duas imagens com alta robustez a variações globais de intensidade nos pixels, e a normalização de seu resultado para um valor entre -1 e 1, resultante da equação acima permite o uso de um limiar que indica a confiabilidade do rastreamento. Novas imagens são incorporadas ao mapa com base em um limiar de 0.97, escolhido empiricamente, garante que a medida de confiança das imagens adicionadas ao mapa seja sempre alta.

O valor da NCC também é usado como indicador de falhas no rastreamento, oriundas de obstruções no campo de visão. Quando a medida de confiança é menor que um limiar de 0.85, o rastreamento é interrompido, um procedimento de redetecção (explicado na Seção 2.1.2) é executado para encontrar a imagem do mapa mais próxima ao campo de visão atual e o mapeamento é reiniciado. A Figura 20 mostra os resultados do método de mapeamento proposto por Richa et al. (2014).

3.1.3 Seleção Eficiente de Pixels

Como já citado, ao contrário de métodos baseados em pontos de interesse, os métodos diretos sofrem de um elevado custo computacional, por usarem uma grande quantidade de informação presente em cada imagem. Alternativas como o uso de placas de vídeo e estratégia de seleção de pixels (BAKER; MATTHEWS, 2002; MEILLAND; COMPORT; RIVES, 2011) são comumente usadas. O *framework* de Richa et al. (2014) usa o método proposto por Meilland, Comport e Rives (2011), devido à restrição de esforço computacional necessária para a execução simultânea ao exame da retina. Além disso, a seleção de pixels de Meilland, Comport e Rives (2011) também considera o modelo de transformação de translação e rotação adotado.

O método busca selecionar os M pixels (uma fração do número total de pixels N) mais adequados para o rastreamento. Primeiramente, para cada canal de cor da imagem, calcula-se uma Jacobiana $J(\mathbf{p}_0)$ de dimensões $N \times 3$. Em seguida, os valores em cada uma das três colunas

de J são organizados em ordem decrescente. Finalmente, a lista de pixels escolhidos é preenchida recursivamente pelas posições dos pixels de cada coluna de J . Essencialmente, o método seleciona os pixels com os gradientes mais elevados no espaço dos parâmetros, ao invés da imagem propriamente dita. Um exemplo do resultado da seleção de pixels é dado na Figura 26.

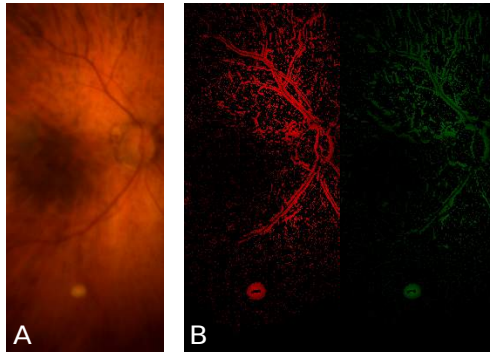


Figura 26 – Resultado do método de seleção eficiente de pixels proposto por Meilland, Comport e Rives (2011). A) Imagem original. B) Imagens resultantes da seleção dos pixels considerados mais representativos para o rastreamento. O número de pixels selecionados corresponde a 20% do número de pixels na imagem original.

3.2 AJUSTE GLOBAL

O método de mapeamento resulta em uma grade de imagens sobrepostas usadas para a criação do mosaico da retina. O limiar de 0.97, aplicado à NCC com o intuito de manter um rastreamento preciso, causa a adição de imagens com grandes regiões em comum a posições distintas do mapa, por vezes correspondendo a mais de 90% da área de duas imagens vizinhas. O acúmulo de erros, no entanto, causa desalinhamentos nas regiões sobrepostas. A junção de duas imagens vizinhas desalinhadas durante a construção de um mapa causa os erros de registro mostrados na Seção 1.1.1.

O método de ajuste global proposto nessa dissertação utiliza a grade de imagens como base, buscando realinhar o mosaico através da análise das regiões de sobreposição nas imagens que o compõe. A estratégia consiste em parear imagens da grade e ajustar os pares si-

multaneamente, propagando as mudanças a partir de uma image fixa no centro até que um estado de convergência global seja atingido. Um modelo de transformação não rígido, limitado por termos de regularização para evitar deformações excessivas, permite que as imagens sejam adequadas ao alinhamento global do mapa. As etapas do processo de ajuste são detalhadas nas seções seguintes.

3.2.1 Pré-processamento

Durante um exame, o operador da lâmpada de fenda regula a posição da fonte de luz para se adequar ao campo de visão atual, o que pode causar grandes variações de luminosidade, mesmo entre imagens vizinhas. Para lidar com isso, modelos de compensação (RICHA et al., 2014) e medidas de similaridade invariantes à iluminação (ASMUTH et al., 2001) são empregados. Estas técnicas, somadas ao alto custo inerente ao uso de métodos diretos, resultam em um aumento de complexidade e esforço computacional do mapeamento.

Para possibilitar um grau viável de robustez a variações de luminosidade, o método proposto usa imagens de gradiente, cujo objetivo é acentuar alterações bruscas de intensidade entre pixels vizinhos. As imagens de gradiente são obtidas através da conversão das imagens RGB de entrada a tons de cinza, seguida da aplicação de um filtro Sobel. Em seguida, com o objetivo de aumentar a região de convergência durante o ajuste, um *kernel* Gaussiano é aplicado:

$$I = I_c * S * G \quad (3.10)$$

onde I_c representa uma imagem de entrada, e S e G representam os operadores Sobel e Gaussiano, respectivamente. O método de obtenção de imagens de gradiente acentua variações de intensidade, o que as permite manter mais informações estruturais relevantes sem um grande aumento no esforço computacional necessário para métodos de compensação de iluminação. A Figura 27 compara imagens de gradiente ao método de compensação de iluminação por aproximação a uma curva TPS (RICHA et al., 2014) e a um filtro Laplaciano da Gaussiana.

3.2.2 Ajuste

O resultado do método de mapeamento consiste em uma matriz cujas posições correspondem ao centro de cada imagem que compõe o

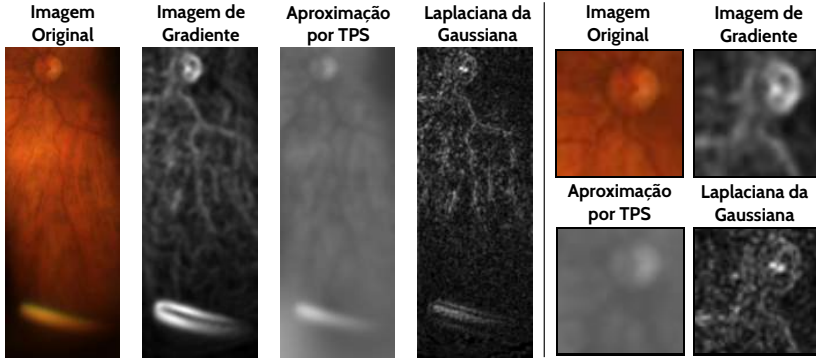


Figura 27 – Comparação de imagens de gradiente, obtidas através da aplicação da Equação 3.10, com o método de compensação de iluminação através da aproximação por uma curva TPS (RICHA et al., 2014) e com um filtro Laplaciano da Gaussiana. Nota-se a acentuação da informação estrutural relevante, como o nervo óptico e vasos sanguíneos na imagem de gradiente.

mosaico da retina. O espaçamento entre imagens vizinhas é de poucos pixels, o que resulta nas grandes regiões de sobreposição que contribuem para a qualidade do mapa. A base para o procedimento de ajuste global é dividir o conjunto de imagens de entrada em uma lista de pares \mathcal{L} e alinhar cada par simultaneamente de acordo com sua região em comum. A construção da lista \mathcal{L} é exemplificada na Figura 28.

Dado um conjunto de imagens de entrada $\mathcal{M} = I_0, I_1, \dots, I_n$, cada par de imagens (I_a, I_b) em \mathcal{L} é alinhado através da minimização de uma medida de erro que representa a similaridade entre suas regiões de sobreposição. I_a e I_b são alteradas por uma função de transformação com vetores de parâmetros \mathbf{p}_a e \mathbf{p}_b , que são atualizados iterativamente até que ocorra a convergência global:

$$\epsilon_d = \sum_{(I_a, I_b) \in \mathcal{L}} \sum_{\mathbf{x}} (I_a(w(\mathbf{x}, \mathbf{p}_a)) - I_b(w(\mathbf{x} - \delta, \mathbf{p}_b)))^2 \quad (3.11)$$

onde $I(w(\mathbf{x}, \mathbf{p}))$ representa o valor da imagem I na posição resultante da aplicação da função w a \mathbf{x} com parâmetros \mathbf{p} . δ representa o deslocamento referente às posições distintas das imagens na grade.

A Equação 3.11 é minimizada partindo de uma estimativa inicial razoável dos parâmetros em cada \mathbf{p} , e calculando um vetor $\Delta \mathbf{p}$ tal que

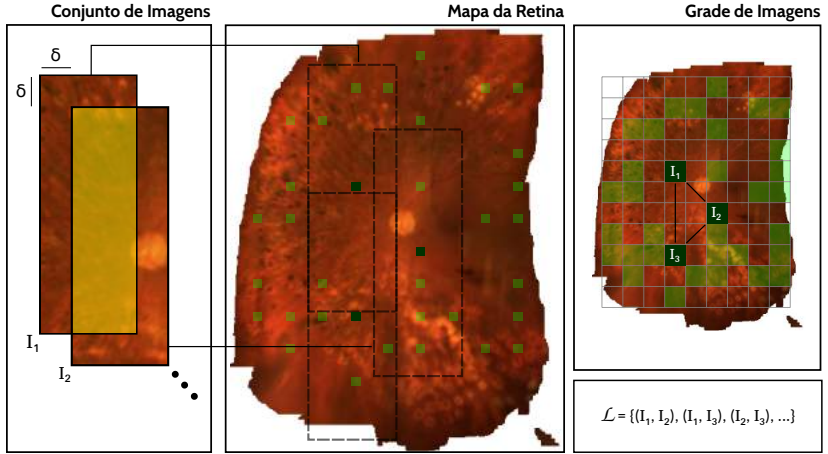


Figura 28 – Processo de construção da lista de pares de imagens usadas como base para o ajuste fino. Esquerda: representação das imagens de entrada, com sua região em comum em destaque. Centro: mapa resultante do método de mapeamento (RICHA et al., 2014), onde cada posição destacada corresponde ao centro de uma das imagens no mosaico. Direita: exemplo de correspondência entre imagens vizinhas para a formação da lista de pares \mathcal{L} .

$\mathbf{p} = \mathbf{p} + \Delta\mathbf{p}$ leva a um valor mínimo de ϵ_d . O valor dos elementos de $\Delta\mathbf{p}$ é obtido segundo o método de minimização de Newton:

$$\Delta\mathbf{p} = -\frac{\frac{\partial\epsilon_d}{\partial\mathbf{p}}}{\frac{\partial^2\epsilon_d}{\partial^2\mathbf{p}}} \quad (3.12)$$

onde $\frac{\partial^2\epsilon_d}{\partial^2\mathbf{p}}$ representa a segunda derivada de ϵ_d com relação a \mathbf{p} . Nota-se que a Equação 3.12 é aplicada a cada imagem contida no mapa e seu vetor de parâmetros de transformação \mathbf{p} correspondente, com exceção da imagem central do mosaico, que é fixa. Dados um conjunto de n imagens de entrada e uma lista \mathcal{L} com m pares de imagens, a primeira derivada de ϵ_d é:

$$\frac{\partial \epsilon_d}{\partial \mathbf{p}} = J = \begin{bmatrix} \frac{\partial \epsilon_d}{\mathbf{p}_0} \\ \frac{\partial \epsilon_d}{\mathbf{p}_1} \\ \cdot \\ \cdot \\ \cdot \\ \frac{\partial \epsilon_d}{\mathbf{p}_n} \end{bmatrix} = \begin{bmatrix} \frac{\partial(\sum_{\mathbf{x}} \Delta I_0(\mathbf{x}))}{\frac{\partial \mathbf{p}_0}} \\ \frac{\partial(\sum_{\mathbf{x}} \Delta I_1(\mathbf{x}))}{\frac{\partial \mathbf{p}_1}} \\ \cdot \\ \cdot \\ \cdot \\ \frac{\partial(\sum_{\mathbf{x}} \Delta I_n(\mathbf{x}))}{\frac{\partial \mathbf{p}_n}} \end{bmatrix} \quad (3.13)$$

onde I_0, I_1, \dots, I_n são as imagens de entrada e ΔI_i representa a diferença da imagem I_i às suas imagens vizinhas, de acordo com as correspondência em \mathcal{L} :

$$\begin{aligned} \Delta I_i &= (I_i(w(\mathbf{x}, \mathbf{p}_a)) - I_j(w(\mathbf{x} - \delta, \mathbf{p}_b)) \\ &\quad + (I_i(w(\mathbf{x}, \mathbf{p}_a)) - I_k(w(\mathbf{x} - \delta, \mathbf{p}_b)) \\ &\quad + \dots \\ &\quad + (I_i(w(\mathbf{x}, \mathbf{p}_a)) - I_l(w(\mathbf{x} - \delta, \mathbf{p}_b))) \end{aligned} \quad (3.14)$$

A primeira derivada de ϵ_d é, portanto:

$$\frac{\partial \epsilon_d}{\partial \mathbf{p}} = 2 \sum_{I_i \in \mathcal{M}} \sum_{\mathbf{x}} \left(\Delta I_i(\mathbf{x}) \bullet \frac{\partial I_i(w(\mathbf{x}, \mathbf{p}_i))}{\partial \mathbf{p}_i} \right) \quad (3.15)$$

onde o operador \bullet representa o produto escalar.

Para evitar complexidade excessiva e limitar o esforço computacional do método, a segunda derivada de ϵ_d é estimada segundo a aproximação de Gauss-Newton, ou seja:

$$\frac{\partial^2 \epsilon_d}{\partial \mathbf{p}} = H \approx \left(\frac{\partial \epsilon_d}{\partial \mathbf{p}} \right)^2 \quad (3.16)$$

Finalmente, substituindo as Equações 3.13 e 3.16 na Equação 3.12, tem-se:

$$\Delta \mathbf{p} = H^{-1} J \quad (3.17)$$

O resultado da Equação 3.17 é aplicado iterativamente ao valor atual de \mathbf{p} , atualizando os parâmetros de transformação correspondentes a cada imagem até que o sistema convirja ou um número máximo de iterações seja atingido.

Um fator importante para o resultado do ajuste é o modelo de transformação aplicado a cada imagem pela função $w(\mathbf{x}, \mathbf{p})$. A quantidade de parâmetros em \mathbf{p} influi diretamente no esforço computacional

do cálculo de $\Delta \mathbf{p}$ (Equação 3.17). Além disso, a escolha de um modelo que não se adequa às possíveis deformações que ocorrem durante o mapeamento pode resultar em divergência durante processo de ajuste. O modelo adotado foi, o proposto por Lim e Yang (2005), baseado na aproximação por *Thin-Plate Splines* (TPS) apresentada por Duchon (1977).

A estratégia de Lim e Yang (2005) consiste em usar um vetor $\mathbf{c} = \{(x_1, y_1), (x_2, y_2), \dots, (x_k, y_k)\}$ de pontos de controle, distribuídos uniformemente em uma imagem. Cada posição da imagem é interpolada por uma função $f(x, y)$, que descreve uma *spline* passando pelos pontos em \mathbf{c} :

$$f(x, y) = a_1 + a_2x + a_3y + \sum_{i=1}^k v_i U(||(x_i, y_i) - (x, y)||) \quad (3.18)$$

onde $U(r) = r^2 \log r^2$. Nota-se que o vetor \mathbf{p} que representa os parâmetros da transformação TPS é formado pelo vetor $\mathbf{a} = \{a_1, a_2, a_3\}$ e pelos pesos aplicados a cada ponto de controle, representados pelo vetor \mathbf{w} , o que gera uma extensão de uma transformação geométrica padrão (rotação, translação e escala). Exemplos da aplicação das TPS no registro de imagens são mostrados na Figura 29.

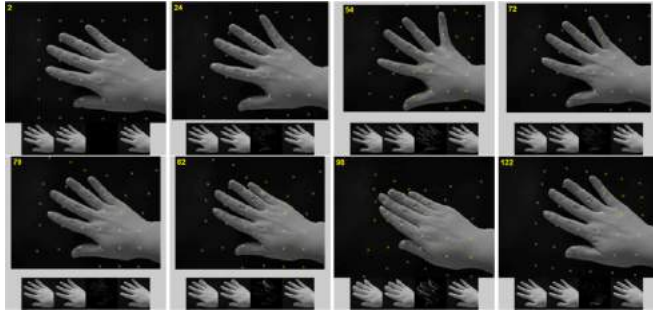


Figura 29 – Exemplo da aplicação de *Thin-Plate Splines* no registro de imagens. Cada imagem ampliada representa a imagem de entrada, que será transformada para se alinhar a uma imagem fixa. Abaixo, a imagem transformada pelo método proposto por Lim e Yang (2005), a diferença entre o resultado e a imagem fixa, e a própria imagem fixa, respectivamente.

O modelo de transformação com base em *Thin-Plate Splines* é

viável no contexto do ajuste global de mapas da retina devido às deformações imprevisíveis sofridas por cada imagem do mosaico. Apesar de o uso da lâmpada de fenda acarretar apenas variações de rotação e translação, a propagação das correções no alinhamento baseado em um modelo de transformação rígido pode causar divergências no ajuste, pois uma única imagem pode possuir diversas vizinhas na grade. No entanto, os resultados mostrados na Figura 29 mostram deformações excessivas no registro de pares de imagens particularmente distintos. Com o objetivo de evitar que estas deturpações afetem o mapa final, os seguintes termos de regularização foram adicionados à medida de erro na Equação 3.11:

$$\epsilon_a = \sum_{(I_a, I_b) \in \mathcal{L}} \sum_h ({}^h \mathbf{p}_a - {}^h \mathbf{p}_b - {}^h \mathbf{d})^2 \quad (3.19)$$

$$\epsilon_b = \sum_{I_a \in \mathcal{M}} \sum_h ({}^h \mathbf{p}_a - {}^h \mathbf{p}_*)^2 \quad (3.20)$$

onde h é o número de parâmetros de \mathbf{p} , cujo acesso em uma dada posição i é representado por ${}^i \mathbf{p}$. \mathbf{d} é um vetor contendo as distâncias relativas entre cada par de imagens em \mathcal{L} , de modo a compensar pelas posições deslocadas na grade. \mathbf{p}_* representa a configuração inicial de \mathbf{p} .

Na Equação 3.19, o termo ϵ_a regulariza deformações entre cada par de imagens, evitando afastamento excessivo entre os respectivos pontos de controle. O termo ϵ_b na Equação 3.20 limita a transformação individual de cada imagem, impedindo que os pontos de controle da TPS se afastem de sua configuração inicial, evitando assim alterações das estruturas do mosaico original. Combinando as Equações 3.11, 3.19 e 3.20, tem-se uma medida de convergência global do processo de ajuste fino:

$$\epsilon_t = [\epsilon_d + \sigma_{ab}(\lambda_a \epsilon_a + \lambda_b \epsilon_b)] \quad (3.21)$$

onde σ_{ab} , λ_a e λ_b são coeficientes usados para controlar o impacto dos termos de regularização; σ_{ab} é uma função de similaridade baseada na correlação cruzada normalizada (NCC) entre cada par de imagens (I_a e I_b), definida como $1 - NCC(I_a, I_b)$, que garante que os termos de regularização dominem a equação em pares cuja similaridade é baixa. λ_a e λ_b são pesos escolhidos empiricamente que regulam os efeitos individuais de ϵ_a e ϵ_b . O ajuste fino consiste, portanto, em encontrar o conjunto de parâmetros $\Delta \mathbf{p}$, de modo a minimizar a medida de similaridade ϵ_t .

3.3 RESULTADOS

Para testar os resultados do método, um banco de dados com 86 vídeos obtidos durante exames realizados com lâmpadas de fenda foi usado. O procedimento de mapeamento foi aplicado a todos os vídeos para a criação dos mosaicos da retina, porém em vídeos de baixa qualidade, resultante de fatores como baixa resolução, pouca luminosidade ou campo de visão limitado, não foi possível gerar mapas viáveis para os testes. Dos 86 vídeos no banco de dados, foram criados 15 mosaicos, aos quais foi aplicado o ajuste global. Todos os mapas foram formados por imagens de dimensões 270x100 pixels, com o espaçamento δ entre cada imagem fixado em 6 pixels. O número de imagens resultante do mapeamento e usado como conjunto de entrada variou de 80 a 300. O tamanho mínimo usado para a filtragem de pares com uma região de sobreposição válida foi 30% da área total de uma imagem. O valor de σ para o *kernel* Gaussiano aplicado durante a criação das imagens de gradiente foi 2.

Durante os experimentos, a tendência à convergência ou à divergência no ajuste era claramente notável antes das primeiras 150 iterações pelo método de minimização por Gauss-Newton. Dentre os casos de convergência, o maior número de iterações atingido foi menor que 300. O número máximo de iterações do algoritmo foi, portanto, fixado em 300. Os pesos λ_a e λ_b , que controlam o impacto dos termos de regularização ϵ_a e ϵ_b foram obtidos empiricamente, buscando garantir a flexibilidade necessária para a transformação não-rígida do modelo TPS, mas evitar deformações excessivas. Os valores de λ_a e λ_b foram fixados em 1 e 3, respectivamente.

O número de pontos de controle usados no modelo de transformação por TPS foi 6, espaçados uniformemente por cada imagem em três linhas e duas colunas. O número limitado de pontos contribui para o controle da deformação das imagens e redução do custo computacional, sem afetar consideravelmente o resultado do alinhamento.

Os vídeos foram gravados com uma lâmpada de fenda modelo CSO SL9800 LED (Figura 30). Todos os experimentos foram realizados em um computador com um processador Inter Core i7-4790, com 16GB de memória RAM. Os parâmetros foram mantidos constantes durante todos os experimentos. A linguagem de programação usada para o desenvolvimento foi C++, com auxílio da biblioteca de processamento de imagens OpenCV.

Os resultados esperados para o ajuste global de imagens da retina são mosaicos de boa qualidade e fiéis à realidade, ou seja, imagens

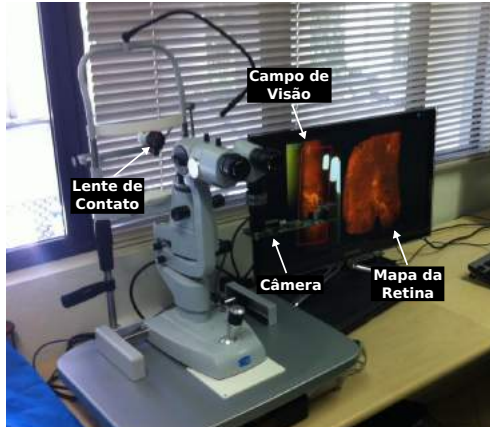


Figura 30 – Lâmpada de fenda CSO SL9800 LED, acoplada a uma câmera conectada a um computador. Na tela, uma demonstração do *framework* de mapeamento da retina proposto por Richa et al. (2014).

claras, nítidas, e sem estruturas desalinhadas. Na Figura 31 são mostrados resultados positivos do procedimento de ajuste. Na esquerda, os mosaicos originais, resultantes da técnica de mapeamento, comparados ao mosaicos ajustados. Uma visão detalhada das regiões de interesse é mostrada na direita. Nota-se claramente que os vasos sanguíneos foram reajustados às suas posições corretas. Degradações como borrões e estruturas duplicadas também foram corrigidas, resultando em mosaicos mais nítidos e bem definidos.

Devido à estratégia incremental usada, no entanto, é possível que o algoritmo atinja um mínimo local durante a minização. O resultado disso é uma convergência errônea do conjunto de imagens. A Figura 32 (A) mostra como isso afeta o mosaico final. Na área destacada, nota-se que o mosaico original possui dois vasos sanguíneos originados no nervo óptico. Durante o ajuste, as imagens contendo essa região tendem a se aproximar, buscando a alta similaridade entre os dois vasos, convergindo para um mínimo incorreto. A estrutura resultante é uma junção das duas originais.

Regiões borradas ou fora de foco e a total ausência de estruturas salientes nas imagens também podem afetar o resultado final do ajuste. Como explicado na Seção 3.2.1, as imagens de gradiente que servem de base para o método destacam regiões onde há variações de intensidade. Imagens contendo grandes regiões uniformes causam di-

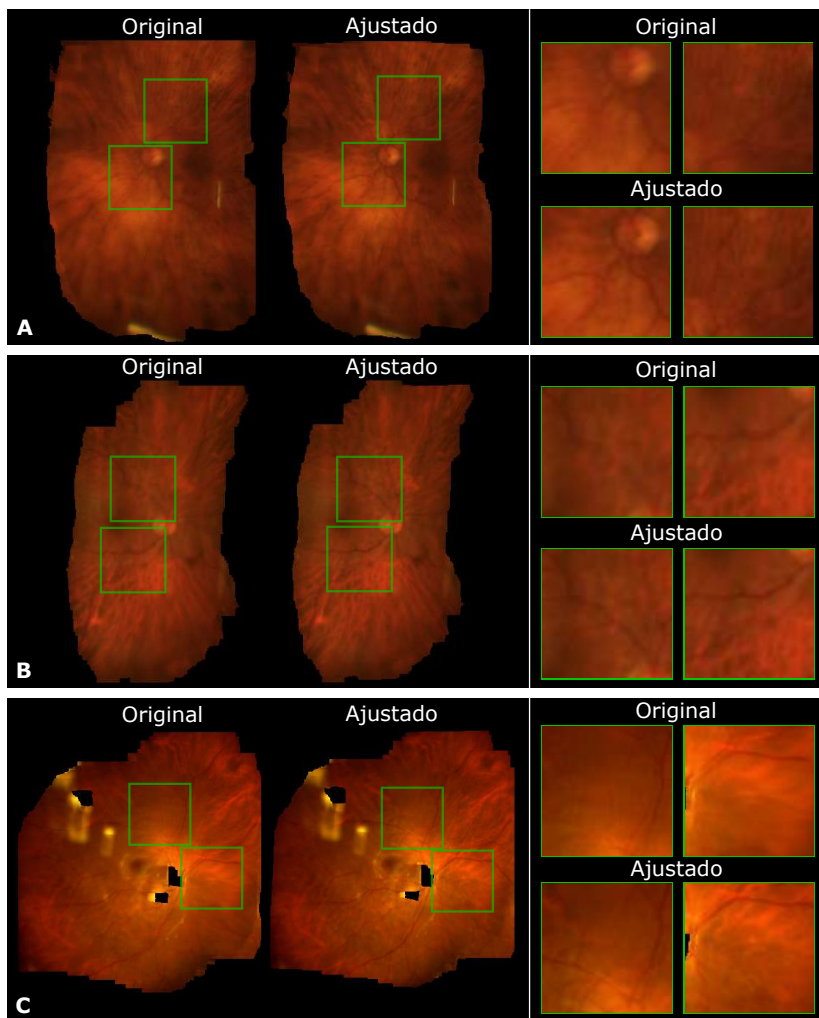


Figura 31 – Esquerda: comparação entre mapas originais, resultantes do método apresentado por Richa et al. (2014) e a estratégia de refinamento proposta. Os resultados mostram um aumento claro na qualidade dos mosaicos, assim como a correção dos erros causados por imprecisões. Regiões de interesse são enfatizadas na direita.

vergências no processo de alinhamento, resultando em mosaicos deformados. Na Figura 32 (B) é mostrado o resultado da tentativa de ajuste

de um mosaico com grandes regiões uniformes (em destaque). Nota-se a irregularidade do resultado devido à divergência das imagens.

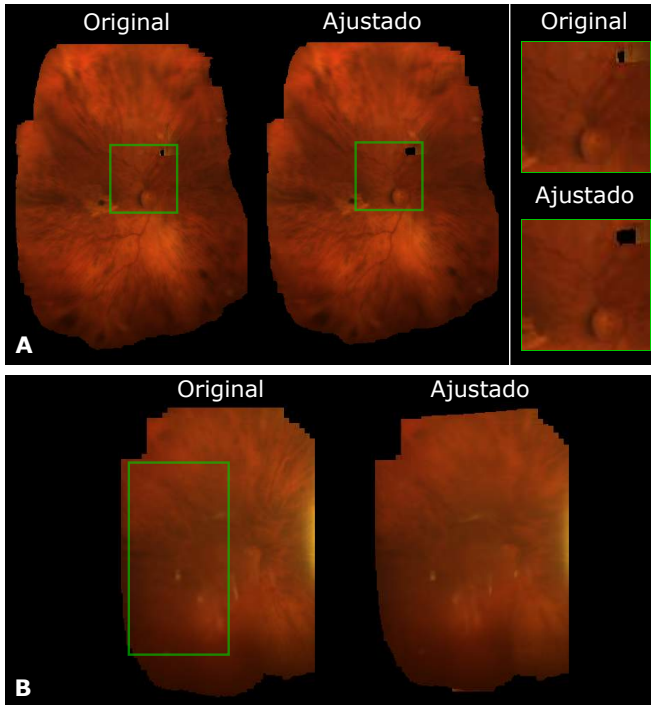


Figura 32 – Exemplos de possíveis erros que podem ocorrer durante o processo de ajuste. Acima: devido à convergência das imagens para um máximo local, ocorre uma junção de dois vasos sanguíneos (em destaque), afetando a qualidade do mosaico final. Abaixo: grandes regiões uniformes no mosaico, como a destacada, impedem a convergência global das imagens para um único mínimo, o que produz um mosaico deformado.

Os resultados visuais mostram que o método proposto de ajuste global permite aumentar a qualidade final de mosaicos da retina, embora haja uma restrição quanto à natureza das imagens de entrada. Um dos obstáculos para a criação de um mapa bem ajustado é a ausência de variações locais de intensidade. A retina é rica em textura, como nota-se na Figura 31, portanto é importante que a obtenção do vídeo de entrada seja feita adequadamente, em foco e com iluminação propriamente manipulada.

Para quantificar o ajuste dos mosaicos, levou-se em consideração que as imagens centrais das grades que os formam permanecem fixas durante o processo, o que causa a convergência ao centro do mosaico. A grande área de sobreposição entre cada par de imagens também permite medir o progresso do ajuste através da correlação cruzada normalizada (NCC). A NCC é medida para cada região de sobreposição presente entre os pares de imagens. À medida que o ajuste é realizado, um aumento global da similaridade dessas regiões indica que os pares estão convergindo corretamente. A Figura 33 demonstra exemplos da variação da NCC durante a execução do ajuste. No início do processo, quando as imagens estão desalinhadas, grandes modificações são aplicadas a cada par, o que causa um aumento brusco na NCC média. À medida em que as imagens são registradas, a similaridade estabiliza antes da convergência.

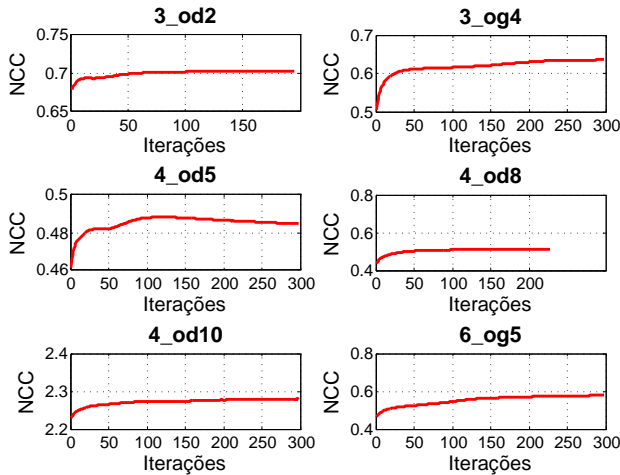


Figura 33 – Aumento da similaridade média em seis mosaicos diferentes, durante o ajuste global. Nota-se que há um aumento brusco inicialmente, quando as mudanças mais representativas são aplicadas aos pares de imagens. A NCC estabiliza ao final da execução, quando a convergência é atingida.

Nos casos em que não há convergência para a imagem central, no entanto, a similaridade global sofre devido às deformações errôneas sofridas por cada imagem. A Figura 34, mostra uma baixa na NCC média à medida que cada imagem se distancia do centro do mosaico.

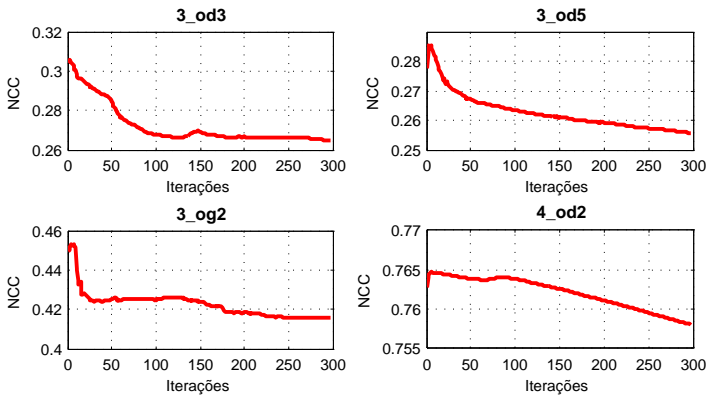


Figura 34 – Baixa da similaridade média em quatro mosaicos durante o ajuste global. Alguns casos apresentam um leve aumento da NCC nas primeiras iterações, que indica uma execução inicial ideal. Quando a divergência passa a ocorrer, no entanto, a similaridade diminui constantemente até o final da execução.

4 CONCLUSÃO

A escassez de pontos-chave é uma das principais limitações para o tratamento das imagens obtidas por microscópios com lâmpada de fenda, tornando inviável o uso de métodos baseados em pontos de interesse, tanto para o mapeamento quanto para o ajuste das imagens. Os resultados do procedimento de ajuste global proposto nessa dissertação mostram o potencial de uma abordagem baseada por métodos diretos para o alinhamento de mosaicos da retina.

Apesar das vantagens oferecidas pela lâmpada de fenda, grande parte do esforço acadêmico voltado ao mapeamento da retina foca em imagens da câmera de fundo. Consequentemente, o suporte a exames que necessitam de um nível de detalhe maior, como a fotocoagulação a laser, é limitado. Mesmo as estratégias de mapeamento existentes sofrem dos erros inerentes ao uso de métodos diretos. O método proposto busca preencher esta lacuna e prover uma nova ferramenta de auxílio para exames e cirurgias realizadas através da lâmpada de fenda.

O método proposto foi testado em vídeos obtidos com a lâmpada de fenda, onde os desafios citados na Seção 1.1.1 são frequentes. Apesar disso, os resultados mostram melhorias visíveis nos mapas, o que atesta a viabilidade da aplicação da estratégia em exames médicos reais. Erros como desalinhamentos, borrões e a presença de *ghosting* são tratados, produzindo assim um mosaico de maior qualidade e fiel à realidade. Embora a estratégia proposta, assim como os métodos baseados em pontos de interesse, dependa da presença de estruturas salientes, o ajuste ainda se mostra possível em situações nas quais o uso de pontos de interesse seria inviável. Apesar das restrições relativas à qualidade das imagens de entrada, o método mostra um grande potencial de uso para a assistência computacional em exames oftalmológicos feitos com a lâmpada de fenda.

O método também é capaz de lidar com os desafios inerentes ao uso de imagens da lâmpada de fenda, que serviram como base para os objetivos específicos dessa dissertação (Seção 1.2). Especificamente, as falhas causadas pelo acúmulo de erros devido a imperfeições no rastreamento são tratadas pela estratégia de ajuste global (Seção 3.2), cujas possíveis deformações excessivas são abordadas através do uso de termos de regularização apresentados na Seção 3.2.2. Por último, eventuais variações de iluminação entre imagens causadas pelo reposicionamento da fonte de luz são tratadas através do uso de imagens de gradiente (Seção 3.2.1).

REFERÊNCIAS

AABERG, T. M. **Proliferative Diabetic Retinopathy (PDR), PRP, Regressed NV**. 2013. Disponível em: <<http://imagebank.asrs.org/file/3535/proliferative-diabetic-retinopathy-pdr-prp-regressed-nv>>.

ASMUTH, J. et al. Mosaicking and enhancement of slit lamp biomicroscopic fundus images. **British Journal of Ophthalmology**, v. 85, n. 5, p. 563–565, 2001.

ATASOY, S. et al. A Global Approach for Automatic Fibroscopic Video Mosaicing in Minimally Invasive Diagnosis. In: (**MICCAI**). [S.l.: s.n.], 2008. p. 850–857.

BAKER, S.; MATTHEWS, I. Equivalence and efficiency of image alignment algorithms. In: **Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2001. CVPR 2001**. [S.l.: s.n.], 2001. v. 1, p. I1090–I1097.

BAKER, S.; MATTHEWS, I. **Lucas-Kanade 20 Years On: A Unifying Framework: Part 1**. Pittsburgh, PA, July 2002.

BATHINA, Y.; MEDATHATI, N.; SIVASWAMY, J. Robust matching of multi-modal retinal images using radon transform-based local descriptor. In: **Proceedings of the 1st ACM International Health Informatics Symposium**. [S.l.: s.n.], 2010. p. 765–770.

BAY, H. et al. Speeded-up robust features (surf). **Computer Vision and Image Understanding**, New York, NY, USA, v. 110, n. 3, p. 346–359, June 2008.

BEHRENS, A. et al. Real-time image composition of bladder mosaics in fluorescence endoscopy. **Computer Science - Research and Development**, v. 26, n. 1-2, p. 51–64, 2010.

BENHIMANE, S.; MALIS, E. Homography-based 2d visual tracking and servoing. **Int. J. Rob. Res.**, Sage Publications, Inc., Thousand Oaks, CA, USA, v. 26, n. 7, p. 661–676, July 2007. ISSN 0278-3649. Disponível em: <<http://dx.doi.org/10.1177/0278364907080252>>.

BERGEN, J. et al. Hierarchical model-based motion estimation. In: SANDINI, G. (Ed.). **Computer Vision — ECCV'92: Second European Conference on Computer Vision Santa Margherita Ligure, Italy, May 19–22, 1992 Proceedings**. Berlin, Heidelberg: Springer Berlin Heidelberg, 1992. p. 237–252.

BLACK, M.; JEPSON, A. Eigentracking: Robust matching and tracking of articulated objects using a view-based representation. **International Journal of Computer Vision**, v. 26, n. 1, p. 63–84, 1998.

BONTALA, A.; SIVASWAMY, J.; PAPPURU, R. Image mosaicing of low quality neonatal retinal images. In: **Proceedings of International Symposium on Biomedical Imaging**. [S.l.: s.n.], 2012. p. 720–724.

BROEHAN, A. et al. Real-time Multimodal Retinal Image Registration for Computed-Assisted Laser Photocoagulation System. **IEEE Transactions on Biomedical Engineering (TBME)**, v. 58, n. 10, p. 2816–2824, 2011.

BURT, P.; ADELSON, E. A multiresolution spline with application to image mosaics. **ACM Transactions on Graphics**, 1983.

CAN, A. et al. A feature-based, robust, hierarchical algorithm for registering pairs of images of the curved human retina. **IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)**, v. 24, n. 3, p. 347–364, 2002.

CAN, A. et al. A feature-based technique for joint, linear estimation of high-order image-to-mosaic transformations: mosaicing the curved human retina. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 24, n. 3, p. 412–419, March 2002. ISSN 0162-8828.

CAN, A. et al. Rapid automated tracing and feature extracton from retina fundus images using direct exploratory algorithms. **IEEE Transactions on Information Technology in Biomedicine**, v. 3, n. 2, p. 125–138, June 1999.

CHANWIMALUANG, T.; FAN, G. An efficient algorithm for extraction of anatomical structures in retinal images. In: **Proceedings of 2003 International Conference on Image Processing, 2003. ICIP 2003**. [S.l.: s.n.], 2003. v. 1, p. I1093–I1096.

CHANWIMALUANG, T.; FAN, G. An efficient blood vessel detection algorithm for retinal images using local entropy thresholding. In: **Proceedings of the 2003 International Symposium on Circuits and Systems, 2003. ISCAS '03.** [S.l.: s.n.], 2003. v. 5, p. V21–V24.

CHANWIMALUANG, T.; FAN, G.; FRANSEN, S. Hybrid retinal image registration. **IEEE Transactions on Information Technology in Biomedicine**, v. 10, n. 1, p. 129–142, 2006.

CHOE, T. et al. Optimal global mosaic generation from retinal images. In: **Proceedings of the International Conference on Pattern Recognition (ICPR'06).** Washington, DC, USA: [s.n.], 2006. p. 681–684.

COOTES, T.; EDWARD, G.; TAYLOR, C. Active appearance models. In: BURKHARDT, H.; NEUMANN, B. (Ed.). **Computer Vision — ECCV'98: 5th European Conference on Computer Vision Freiburg, Germany, June 2–6, 1998 Proceedings, Volume II.** Berlin, Heidelberg: Springer Berlin Heidelberg, 1998. p. 484–498.

DAME, A.; MARCHAND, E. Second order optimization of mutual information for real-time image registration. **IEEE Trans. on Image Processing**, v. 21, n. 9, p. 4190–4203, 2012.

DUCHON, J. Splines minimizing rotation-invariant semi-norms in sobolev spaces. In: _____. **Constructive Theory of Functions of Several Variables: Proceedings of a Conference Held at Oberwolfach April 25 – May 1, 1976.** Berlin, Heidelberg: Springer Berlin Heidelberg, 1977. p. 85–100. ISBN 978-3-540-37496-1. Disponível em: <<http://dx.doi.org/10.1007/BFb0086566>>.

FLORACK, L. et al. General intensity transformations and differential invariants. **Journal of Mathematical Imaging and Vision**, v. 4, n. 2, p. 171–187, 1994.

HAGER, G.; ; BELHUMEUR, P. Efficient region tracking with parametric models of geometry and illumination. **IEEE Trans. Pattern Anal. Mach. Intell.**, IEEE Computer Society, Washington, DC, USA, v. 20, n. 10, p. 1025–1039, out. 1998. ISSN 0162-8828. Disponível em: <<http://dx.doi.org/10.1109/34.722606>>.

HARRIS, C.; STEPHENS, M. A combined corner and edge detector. In: **Proceedings of Fourth Alvey Vision Conference**. [S.l.: s.n.], 1988. p. 147–151.

KALAL, Z.; MATAS, J.; MIKOLAJCZYK, K. P-N learning: Bootstrapping binary classifiers by structural constraints. In: **(CVPR)**. [S.l.: s.n.], 2010. p. 49–56.

KAPLAN, H. J.; PIRI, N. **PRP Laser**. 2013. Disponível em: <<http://imagebank.asrs.org/file/5338/prp-laser>>.

KITCHENHAM, B.; CHARTERS, S. **Guidelines for performing Systematic Literature Reviews in Software Engineering**. [S.l.], 2007.

LIM, J.; YANG, M. A direct method for modeling non-rigid motion with thin plate spline. In: **Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR'05)**. Washington, USA: [s.n.], 2005. p. 1196–1202.

LINDENBERG, T. Feature detection with automatic scale selection. **International Journal of Computer Vision**, v. 30, p. 79–116, 1998.

LINHARES, R. et al. Local binary pattern matching for fast retina map relocalization using the slit-lamp. In: HANCOCK, E. B.-C. E. (Ed.). **Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications: 19th Iberoamerican Congress, CIARP 2014, Puerto Vallarta, Mexico, November 2-5, 2014. Proceedings**. [S.l.]: Springer International Publishing, 2014. p. 478–485.

LINHARES, R. et al. Non-rigid fine adjustment of retina maps acquired using a slit-lamp. In: **2015 IEEE 28th International Symposium on Computer-Based Medical Systems, CBMS 2016**. Dublin e Belfast, Irlanda: [s.n.], 2016.

LOVEGROVE, S. **Parametric Dense Visual Slam**. Tese (Doutorado) — Department of Computing, Imperial College London, London, UK, 2011.

LOWE, D. Distinctive image features from scale-invariant keypoints. **International Journal of Computer Vision**, v. 60, n. 2, p. 91–110, 2004.

LUCAS, B.; KANADE, T. An Iterative Image Registration Technique with an Application to Stereo Vision (IJCAI). In: **Proceedings of the 7th International Joint Conference on Artificial Intelligence (IJCAI '81)**. [S.l.: s.n.], 1981. p. 674–679.

MEILLAND, M.; COMPORT, A. I.; RIVES, P. Dense visual mapping of large scale environments for real-time localisation. In: **IEEE International Conference on Intelligent Robots and Systems**. [S.l.: s.n.], 2011. p. 4242–4248.

National Institute of Health (NIH). National eye institute clinical studies database: Early treatment diabetic retinopathy study (etdrs). In: . [s.n.], 2014. Disponível em: <www.nei.nih.gov/neitrials/static/study53.asp>.

RICHA, R. et al. Fundus image mosaicking for information augmentation in computer-assisted slit-lamp imaging. **IEEE Transactions on Medical Imaging (TMI)**, 2014.

RICHA, R.; SZNITMAN, R.; HAGER, G. **Robust Similarity Measures for Gradient-based Direct Visual Tracking**. [S.l.], 2012. Disponível em: <www.sites.google.com/site/rogerioricha/publications>.

RICHA, R. et al. Hybrid Tracking and Mosaicking for Information Augmentation in Retinal Surgery. In: **(MICCAI)**. Nice, France: [s.n.], 2012. p. 397–404.

ROSA, B. et al. Building Large Mosaics of Confocal Endomicroscopic Images Using Visual Servoing. **IEEE Transactions on Biomedical Engineering**, v. 60, n. 4, p. 1041–1049, 2013.

SILVEIRA, G.; MALIS, E. Unified direct visual tracking of rigid and deformable surfaces under generic illumination changes in grayscale and color images. In: . [S.l.]: Springer US, 2010. v. 89, n. 1, p. 84–105.

TOTZ, J. et al. Dense Surface Reconstruction for Enhanced Navigation in MIS. In: **(MICCAI)**. Toronto, Canada: Springer, 2011. (Lecture Notes in Computer Science (LNCS)), p. 89–96.

World Health Organization. **Global Data on Visual Impairments 2010**. [S.l.], 2010. Disponível em: <<http://www.who.int/blindness/GLOBALDATAFINALforweb.pdf>>.

YE, M. et al. Pathological Site Retargeting under Tissue Deformation Using Geometrical Association and Tracking. In: (**MICCAI**). [S.l.]: Springer, 2013, (Lecture Notes in Computer Science (LNCS), v. 8150). p. 67–74.